

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
30 September 2004 (30.09.2004)

PCT

(10) International Publication Number
WO 2004/083371 A2

- (51) International Patent Classification⁷: C12N (74) Agent: ASTRAZENECA; Global Intellectual Property, P
O Box 272, Mereside, Macclesfield, Cheshire SK10 4GR
(GB).
- (21) International Application Number: PCT/GB2004/001124
- (22) International Filing Date: 16 March 2004 (16.03.2004)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
0306185.0 19 March 2003 (19.03.2003) GB
- (71) Applicant (for AE, AG, AL, AM, AT, AU, AZ, BA, BB, BE, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, SZ, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW only): ASTRAZENECA AB [SE/SE]; S-151 85 Sodertälje (SE).
- (71) Applicant (for MG only): ASTRAZENECA UK LIMITED [GB/GB]; 15 Stanhope Gate, London, Greater London W1K 1LN (GB).
- (72) Inventor; and
- (75) Inventor/Applicant (for US only): DAHL, Niklas [SE/SE]; Uppsala University, Årsgatan 5, S-752 36 Uppsala (SE).
- (81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:
— without international search report and to be republished upon receipt of that report
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: MOLECULES

(57) Abstract: The present invention describes the chromosomal location of obesity susceptibility genes and provides methods for detecting the presence of such genes. The present invention enables the development of novel therapies to treat obesity.

WO 2004/083371 A2

JPATS
JC05 Rec'd PCT/PTO 16 SEP 2005

- 1 -

MOLECULES

The present invention describes the chromosomal location of one or more genes involved in obesity.

5 In recent years, there has been an increase in the population of patients with what are termed "life-style related" diseases, such as obesity and hyperlipidemia. Obesity refers to a physical state in which the stored fat, constituted mainly of triglycerides, is accumulated excessively in the body. A consequence of obesity is an increased risk of arteriosclerosis, fatty liver, hypertension, diabetes, coronary heart disease, stroke, gallbladder disease,
10 osteoarthritis, respiratory problems and some types of cancer (Visscher & Seidell, Annu Rev Public Health (2001); 22(1): 355-75). It has been shown recently that there can also exist a strong genetic component underlying the development of obesity. The role of certain genes, e.g., *agouti*, leptin (*ob*) and leptin receptor (*db*), in the regulation of body fat distribution has been described (Friedman & Halaas, Nature (Oct 1998); 395(6704): 763-70; Ahima et al,
15 Front Neuroendocrinol (Jul 2000); 21(3): 263-307). In humans, mutations in leptin, leptin receptor, melanocortin 4- receptor (MC4-R) and peroxisome proliferator-activated receptor (PPAR) γ 2 genes have been shown to be important component in patients with severe obesity (Clement *et. al.*; Int J Obes Relat Metab Disord 2000 Mar ;24(3): 391-3, Yeo *et. al.*; Nat Genet 1998 Oct;20(2): 111-2).

20 The identification of other genes involved in the development of obesity will provide a better understanding of the disease, and an opportunity to develop more specific and effective drugs.

The present invention is based on the discovery of a chromosomal translocation which has a predisposing effect on the development of obesity. The translocation exists between
25 chromosome 4 and chromosome 15. Specifically, chromosome 4 has been broken at cytoband 4q22.3 and chromosome 15 has been broken at cytoband 15q22.2. The outcome of the translocation event is that at least two genes are affected in persons having this translocation and that persons having this translocation are more susceptible to develop obesity.

30 Two of the genes affected by the translocation are uncoordinated 5C (UNC5C) residing on chromosome 4 and RAR-like orphan receptor alpha 1 (ROR α 1) residing on chromosome 15.

- 2 -

The breakpoint at chromosome 4 occurs between the genes pyruvate dehydrogenase A2 (PDHA2) and UNC5C. While neither of these genes are destroyed in the mutation event, UNC5C transcription is upregulated in adipocytes. A cDNA encoding UNC5C is set out as SEQ ID NO: 9 (with the first nucleotide of the UNC5C coding region accorded position 155),
5 and its gene product is set out as SEQ ID NO: 10.

There is an alternative splice version produced from the UNC5C gene, where the exons 11 and 12 are spliced together from internal sites producing a transcript with a truncated open reading frame and a much shorter protein. A sequence containing this splice event has been determined and submitted to the GENESEQN database and has accession
10 number AAC90914 (Fig. 2 and Fig.8). We confirmed the presence of this alternative transcript with the use of PCR on human brain material. A putative "in silico" generated full length clone of the alternative transcript is set out as SEQ ID NO: 13 (with the first nucleotide of the sUNC5C coding region accorded position 1) and its gene product is set out as SEQ ID NO 14. As used herein "sUNC5C" refers to this shorter splice version of UNC5C unless
15 otherwise stated. The sequence AAC90914 is a truncated version of sUNC5C.

The breakpoint at chromosome 15 is within the first intron of ROR α 1 and this mutation event destroys the transcription of the ROR α 1 isoform. A cDNA encoding ROR α 1 is set out as SEQ ID NO: 5 (with the first nucleotide of the ROR α 1 coding region accorded position 102), and its gene product is set out as SEQ ID NO:6. The translocation results in the
20 first exon of ROR α 1 being spliced into the second exon of UNC5C, resulting in the expression of two novel transcripts which are fusions between ROR α 1 and UNC5C (long and short). These two novel transcripts are expressed in affected patients adipocytes. The first 657 amino acids are identical for both fusion proteins.

The first transcript results in the translation of a long fusion protein between ROR α 1-
25 UNC5C (termed herein "IROR α 1-UNC5C"). A cDNA encoding IROR α 1-UNC5C is set out as SEQ ID NO: 1, and its protein fusion product as SEQ ID NO:2. In SEQ ID NO: 1, nucleotide positions 1-233 represent that part of the transcript from ROR α 1. Nucleotide positions 234-2986 represent that part of the transcript from UNC5C. In SEQ ID NO: 2, amino acid positions 1-55 represent that part of the protein fusion product derived from
30 ROR α 1 DNA. Amino acid positions 56-383 and 403-964 represent that part of the protein fusion product derived from UNC5C DNA. The IROR α 1-UNC5C transcript is produced by splicing ROR α 1 exon1 with UNC5C exon 2-16 together with a fusion specific exon situated

- 3 -

between UNC5C exons 7 and 8. The fusion specific exon is an extra exon of 57bp which encodes for an extra 19 amino acid residues and resides within the lROR α 1-UNC5C transcript at nucleotide positions 1220-1276 of SEQ ID NO:1. In SEQ ID NO:2, amino acid positions 384-402 represent that part of the protein fusion product derived from the fusion specific exon. It is to be noted that the first coding nucleotide for the amino acid position 384 comes from UNC5C exon 7.

The second transcript results in the translation of a shorter fusion protein between ROR α 1-UNC5C (known herein as the "sROR α 1-UNC5C"). A cDNA encoding sROR α 1-UNC5C is set out as SEQ ID NO: 3, and its protein fusion product as SEQ ID NO: 4. In SEQ ID NO: 3, nucleotide positions 1-233 represent that part of the transcript from ROR α 1. Nucleotide positions 234-2546 represent that part of the transcript from UNC5C. In SEQ ID NO:4, amino acid positions 1-55 represent that part of the protein fusion product derived from ROR α 1 DNA. Amino acid positions 56-383 and 403-669 represent that part of the protein fusion product derived from sUNC5C DNA. As for the long form, a fusion specific exon is situated between UNC5C exons 7 and 8. The fusion specific exon is an extra exon of 57bp which encodes for an extra 19 amino acid residues and resides within the sROR α 1-UNC5C transcript at nucleotide positions 1220-1276 of SEQ ID NO:3. In SEQ ID NO:4, amino acid positions 384-402 represent that part of the protein fusion product derived from the fusion specific exon. It is to be noted that the first coding nucleotide for the amino acid position 384 comes from UNC5C exon 7. The sROR α 1-UNC5C transcript only differs from lROR α 1-UNC5C in the way that UNC5C exons 11 and 12 are spliced together. sROR α 1-UNC5C uses internal splice sites within UNC5C exon 11 and 12 so that the 3' part of exon 11 and the 5' part of exon 12 is omitted (Fig. 4 and Tables 1 and 2). Thus, sROR α 1-UNC5C has a destroyed open reading frame resulting in the deletion of 307 amino acids and replacing them with 12 amino acids from another reading frame.

Table 1 below shows DNA sequence alignment of the lROR α 1-UNC5C (SEQ ID NO: 1) and sROR α 1-UNC5C (SEQ ID NO: 3) fusion transcripts together with UNC5C(AF055634) and the alternative spliceform of UNC5C (AAC90914). Also shown are the genomic exon sequences of ROR α 1 exon 1 and UNC5C exon 1-16:

- 4 -

Table1.

AF055634 UNC5C_exon1	1 42 CTGCCCTTTGGAGAAAGTGGAGTGTGGCGCTTGGTTGTCGTTAT CTGCCCTTTGGAGAAAGTGGAGTGTGGCGCTTGGTTGTCGTTAT
AF055634 UNC5C_exon1 SEQ ID NO:1 SEQ ID NO:3 RORA1_exon1	43 102 TTCTTCGGACTGCTTCGCGGTGCACGGATTTCAGCTTCTGCCAGTGGGGCTTTCAGCTGT TTCTTCGGACTGCTTCGCGGTGCACGGATTTCAGCTTCTGCCAGTGGGGCTTTCAGCTGT CTGAAAACAGAAGATAGAGGGAGTCTCGGAGCTCGCCATCTCCAGCGATCTCTACATTG CTGAAAACAGAAGATAGAGGGAGTCTCGGAGCTCGCCATCTCCAGCGATCTCTACATTG CTGAAAACAGAAGATAGAGGGAGTCTCGGAGCTCGCCATCTCCAGCGATCTCTACATTG
AF055634 UNC5C_exon1 SEQ ID NO:1 SEQ ID NO:3 RORA1_exon1	103 162 TTGCGCGTCTCTCTGTCCCCCTCCCCCTCCCCCGGCACACCTCTGTCTACGATGAGGAAA TTGCGCGTCTCTCTGTCCCCCTCCCCCTCCCCCGGCACACCTCTGTCTACGATGAGGAAA GGAAAAACATGGAGTTCAGCTCCGGCAGCCCCGACCCCGCCGAGCCAGCCAGCCAGC GGAAAAACATGGAGTTCAGCTCCGGCAGCCCCGACCCCGCCGAGCCAGCCAGCCAGC GGAAAAACATGGAGTTCAGCTCCGGCAGCCCCGACCCCGCCGAGCCAGCCAGCCAGC
AF055634 UNC5C_exon1 SEQ ID NO:1 SEQ ID NO:3 RORA1_exon1	163 222 GGTCTGCGGGCGACAGCGGCCCGCTGCGGACTGGGACTGGGATCTGTCTGCAAAATGCTC GGTCTGCGGGCGACAGCGGCCCGCTGCGGACTGGGACTGGGATCTGTCTGCAAAATGCTC AGCGGCGCGGACGCGGCCCGCCGCTCCAGGGAGACCCGCTGAACCCAGGAATCCGCCCGC AGCGGCGCGGACGCGGCCCGCCGCTCCAGGGAGACCCGCTGAACCCAGGAATCCGCCCGC AGCGGCGCGGACGCGGCCCGCCGCTCCAGGGAGACCCGCTGAACCCAGGAATCCGCCCGC
AF055634 UNC5C_exon1 UNC5C_exon2 SEQ ID NO:1 SEQ ID NO:3 RORA1_exon1	223 282 GTGCTACCTGCCCCGCTGCTCAGCGCCAGCGGCACTGGCTCCGCCGCCCAAGATGAT GTGCTACCTGCCCCGCTGCTCAGCGCCAGCGGCACTGGCTCCGCCGCCCAAG ATGAT AAGAGCGAGCCGCTGCCCCGGTGCAGCAGACAGAGCTATTCCAGCACCAGCAGAGATGAT AAGAGCGAGCCGCTGCCCCGGTGCAGCAGACAGAGCTATTCCAGCACCAGCAGAGATGAT AAGAGCGAGCCGCTGCCCCGGTGCAGCAGACAGAGCTATTCCAGCACCAGCAGAG *****
AF055634 UNC5C_exon2 SEQ ID NO:1 SEQ ID NO:3	283 342 GACTTTTTCATGAACCTCCAGAACTTTTCCTTCTGATCCACCTGAGCCTCTGCCACAT GACTTTTTCATGAACCTCCAGAACTTTTCCTTCTGATCCACCTGAGCCTCTGCCACAT GACTTTTTCATGAACCTCCAGAACTTTTCCTTCTGATCCACCTGAGCCTCTGCCACAT GACTTTTTCATGAACCTCCAGAACTTTTCCTTCTGATCCACCTGAGCCTCTGCCACAT *****
AF055634 UNC5C_exon2 SEQ ID NO:1 SEQ ID NO:3	343 402 TTCTTTATTGAGCCTGAAGAAGCTTATATTGTGAAGAATAAGCCTGTGAACCTGTACTGT TTCTTTATTGAGCCTGAAGAAGCTTATATTGTGAAGAATAAGCCTGTGAACCTGTACTGT TTCTTTATTGAGCCTGAAGAAGCTTATATTGTGAAGAATAAGCCTGTGAACCTGTACTGT TTCTTTATTGAGCCTGAAGAAGCTTATATTGTGAAGAATAAGCCTGTGAACCTGTACTGT *****
AF055634 UNC5C_exon2 SEQ ID NO:1 SEQ ID NO:3	403 462 AAAGCAAGCCCTGCCACCCAGATCTATTTCAGTGTAATAGTGAATGGGTTCATCAGAAG AAAGCAAGCCCTGCCACCCAGATCTATTTCAGTGTAATAGTGAATGGGTTCATCAGAAG AAAGCAAGCCCTGCCACCCAGATCTATTTCAGTGTAATAGTGAATGGGTTCATCAGAAG AAAGCAAGCCCTGCCACCCAGATCTATTTCAGTGTAATAGTGAATGGGTTCATCAGAAG *****
AF055634 UNC5C_exon2 UNC5C_exon3 SEQ ID NO:1 SEQ ID NO:3	463 522 GACCACATAGTAGATGAAAGAGTAGATGAAACTTCCGGTCTCATTTGTCGGGAAGTGAGC GACCACATAGTAGATGAAAGAGTAGATGAAACTTCCG GTCTCATTTGTCGGGAAGTGAGC GACCACATAGTAGATGAAAGAGTAGATGAAACTTCCGGTCTCATTTGTCGGGAAGTGAGC GACCACATAGTAGATGAAAGAGTAGATGAAACTTCCGGTCTCATTTGTCGGGAAGTGAGC *****
AF055634 UNC5C_exon3 SEQ ID NO:1 SEQ ID NO:3	523 582 ATTGAGATTTCGCGCCAGCAAGTGAAGAAGTCTTTGGACCTGAAGATTACTGGTGCCAG ATTGAGATTTCGCGCCAGCAAGTGAAGAAGTCTTTGGACCTGAAGATTACTGGTGCCAG ATTGAGATTTCGCGCCAGCAAGTGAAGAAGTCTTTGGACCTGAAGATTACTGGTGCCAG ATTGAGATTTCGCGCCAGCAAGTGAAGAAGTCTTTGGACCTGAAGATTACTGGTGCCAG *****
AF055634	583 642 TGTGTGGCCTGGAGCTCCGCGGTACCACAAAGAGCCGGAAGGCGTATGTGCGCATTTGCA

- 5 -

UNC5C_exon3
SEQ ID NO:1
SEQ ID NO:3

643 702
TATCTACGGAAGACATTTGAGCAGGAACCCCTAGGAAAGGAAGTGTCTTTGGAACAGGAA
T
ATCTACGGAAGACATTTGAGCAGGAACCCCTAGGAAAGGAAGTGTCTTTGGAACAGGAA
TATCTACGGAAGACATTTGAGCAGGAACCCCTAGGAAAGGAAGTGTCTTTGGAACAGGAA
TATCTACGGAAGACATTTGAGCAGGAACCCCTAGGAAAGGAAGTGTCTTTGGAACAGGAA

AF055634
UNC5C_exon3
UNC5C_exon4
SEQ ID NO:1
SEQ ID NO:3

703 762
GTCCTTACTCCAGTGTGACCCACCTGAAGGGATCCCAGTGGCTGAGGTGGAATGGTTGAAA
GTCCTTACTCCAGTGTGACCCACCTGAAGGGATCCCAGTGGCTGAG GTGGAATGGTTGAAA
GTCCTTACTCCAGTGTGACCCACCTGAAGGGATCCCAGTGGCTGAGGTGGAATGGTTGAAA
GTCCTTACTCCAGTGTGACCCACCTGAAGGGATCCCAGTGGCTGAGGTGGAATGGTTGAAA

AF055634
UNC5C_exon4
UNC5C_exon5
SEQ ID NO:1
SEQ ID NO:3

763 822
AATGAAGACATAATTGATCCCGTTGAAGATCGGAATTTTATATTACTATTGATCACAAC
AATGAAGACATAATTGATCCCGTTGAAGATCGGAATTTTATATTACTATTGATCACAAC
AATGAAGACATAATTGATCCCGTTGAAGATCGGAATTTTATATTACTATTGATCACAAC
AATGAAGACATAATTGATCCCGTTGAAGATCGGAATTTTATATTACTATTGATCACAAC

AF055634
UNC5C_exon5
SEQ ID NO:1
SEQ ID NO:3

823 882
CTCATCATAAAGCAGGCCCGACTCTCTGATACTGCAAAATTACACCTGTGTTGCCAAAAC
CTCATCATAAAGCAGGCCCGACTCTCTGATACTGCAAAATTACACCTGTGTTGCCAAAAC
CTCATCATAAAGCAGGCCCGACTCTCTGATACTGCAAAATTACACCTGTGTTGCCAAAAC
CTCATCATAAAGCAGGCCCGACTCTCTGATACTGCAAAATTACACCTGTGTTGCCAAAAC

AF055634
UNC5C_exon5
UNC5C_exon6
SEQ ID NO:1
SEQ ID NO:3

883 942
ATTGTTGCCAAGAGGAAAAGTACAACCTGCCACTGTCTATAGTCTATGTCAACGGTGGCTGG
ATTGTTGCCAAGAGGAAAAGTACAACCTGCCACTGTCTATAGTCTATG TCAACGGTGGCTGG
ATTGTTGCCAAGAGGAAAAGTACAACCTGCCACTGTCTATAGTCTATG TCAACGGTGGCTGG
ATTGTTGCCAAGAGGAAAAGTACAACCTGCCACTGTCTATAGTCTATG TCAACGGTGGCTGG

AF055634
UNC5C_exon6
SEQ ID NO:1
SEQ ID NO:3

943 1002
TCCACCTGGACGAGTGGTCTGTGTGTAACAGCCGCTGTGGACGAGGGTATCAGAAAACGT
TCCACCTGGACGAGTGGTCTGTGTGTAACAGCCGCTGTGGACGAGGGTATCAGAAAACGT
TCCACCTGGACGAGTGGTCTGTGTGTAACAGCCGCTGTGGACGAGGGTATCAGAAAACGT
TCCACCTGGACGAGTGGTCTGTGTGTAACAGCCGCTGTGGACGAGGGTATCAGAAAACGT

AF055634
UNC5C_exon6
SEQ ID NO:1
SEQ ID NO:3

1003 1062
ACAAGGACTTGTACCAACCCGGCACCCTCAATGGGGGTGCCTTCTGTGAAGGGCAGAGT
ACAAGGACTTGTACCAACCCGGCACCCTCAATGGGGGTGCCTTCTGTGAAGGGCAGAGT
ACAAGGACTTGTACCAACCCGGCACCCTCAATGGGGGTGCCTTCTGTGAAGGGCAGAGT
ACAAGGACTTGTACCAACCCGGCACCCTCAATGGGGGTGCCTTCTGTGAAGGGCAGAGT

AF055634
UNC5C_exon6
UNC5C_exon7
SEQ ID NO:1
SEQ ID NO:3

1063 1122
GTGCAGAAAATAGCCTGTACTACGTTATGCCAGTGGATGGCAGGTGGACGCCATGGAGC
GTGCAGAAAATAGCCTGTACTACGTTATGCCAGTGGATGGCAGGTGGACGCCATGGAGC
TGGATGGCAGGTGGACGCCATGGAGC
GTGCAGAAAATAGCCTGTACTACGTTATGCCAGTGGATGGCAGGTGGACGCCATGGAGC
GTGCAGAAAATAGCCTGTACTACGTTATGCCAGTGGATGGCAGGTGGACGCCATGGAGC

AF055634
UNC5C_exon7
SEQ ID NO:1
SEQ ID NO:3

1123 1182
AAGTGGTCTACTTGTGGAAGTGAAGTGCACCCACTGGCGCAGGAGGGAGTGACCGCGCCA
AAGTGGTCTACTTGTGGAAGTGAAGTGCACCCACTGGCGCAGGAGGGAGTGACCGCGCCA
AAGTGGTCTACTTGTGGAAGTGAAGTGCACCCACTGGCGCAGGAGGGAGTGACCGCGCCA
AAGTGGTCTACTTGTGGAAGTGAAGTGCACCCACTGGCGCAGGAGGGAGTGACCGCGCCA

1183 1242

- 6 -

AF055634
UNC5C_exon7
SEQ ID NO:1
SEQ ID NO:3

GCCCCAAGAATGGAGGCAAGGACTGCGACGGCCTCGTCTTGCAATCCAAGAAGTGCAC
GCCCCAAGAATGGAGGCAAGGACTGCGACGGCCTCGTCTTGCAATCCAAGAAGTGCAC
GCCCCAAGAATGGAGGCAAGGACTGCGACGGCCTCGTCTTGCAATCCAAGAAGTGCAC
GCCCCAAGAATGGAGGCAAGGACTGCGACGGCCTCGTCTTGCAATCCAAGAAGTGCAC

1243 1302
AF055634
UNC5C_exon7
Fusion spec exon
SEQ ID NO:1
SEQ ID NO:3

GATGGGCTTTGCATGCAGA-----
GATGGGCTTTGCATGCAGA
GTTTCATTTATCCCATTTCAACTGAACAGAGAACCCAGAA
GATGGGCTTTGCATGCAGAGTTTCATTTATCCCATTTCAACTGAACAGAGAACCCAGAA
GATGGGCTTTGCATGCAGAGTTTCATTTATCCCATTTCAACTGAACAGAGAACCCAGAA

1303 1362
AF055634
Fusion spec exon
UNC5C_exon8
SEQ ID NO:1
SEQ ID NO:3

-----CTGCTCCTGATTGATGATGTTGCTCTCTATGTTGGGATTGTG
GAATATGGATTTCTT
CTGCTCCTGATTGATGATGTTGCTCTCTATGTTGGGATTGTG
GAATATGGATTTCTTCTGCTCCTGATTGATGATGTTGCTCTCTATGTTGGGATTGTG
GAATATGGATTTCTTCTGCTCCTGATTGATGATGTTGCTCTCTATGTTGGGATTGTG

1363 1422
AF055634
UNC5C_exon8
SEQ ID NO:1
SEQ ID NO:3
AAC90914

ATAGCAGTGATCGTTTGCTGGCGATCTCTGTAGTTGTGGCCTTGTTTGTGTATCGGAAG
ATAGCAGTGATCGTTTGCTGGCGATCTCTGTAGTTGTGGCCTTGTTTGTGTATCGGAAG
ATAGCAGTGATCGTTTGCTGGCGATCTCTGTAGTTGTGGCCTTGTTTGTGTATCGGAAG
ATAGCAGTGATCGTTTGCTGGCGATCTCTGTAGTTGTGGCCTTGTTTGTGTATCGGAAG
TTGTTTGTGTATCGGAAG

1423 1482
AF055634
UNC5C_exon8
SEQ ID NO:1
SEQ ID NO:3
AAC90914

AATCATCGTGACTTTGAGTCAGATATTATGACTCTTCGGCACTCAATGGGGGCTTTTCAN
AATCATCGTGACTTTGAGTCAGATATTATGACTCTTCGGCACTCAATGGGGGCTTTTCAN
AATCATCGTGACTTTGAGTCAGATATTATGACTCTTCGGCACTCAATGGGGGCTTTTCAN
AATCATCGTGACTTTGAGTCAGATATTATGACTCTTCGGCACTCAATGGGGGCTTTTCAN
AATCATCGTGACTTTGAGTCAGATATTATGACTCTTCGGCACTCAATGGGGGCTTTTCAN
AATCATCGTGACTTTGAGTCAGATATTATGACTCTTCGGCACTCAATGGGGGCTTTTCAN

1483 1542
AF055634
UNC5C_exon8
UNC5C_exon9
SEQ ID NO:1
SEQ ID NO:3
AAC90914

CCTGTGAACATCAAGGCAGCAAGACAAGATCTGCTGGCTGTACCCCAAGCTCACGTCA
CCTGTGAACATCAAGGCAGCAAGACAAG
ATCTGCTGGCTGTACCCCAAGCTCACGTCA
CCTGTGAACATCAAGGCAGCAAGACAAGATCTGCTGGCTGTACCCCAAGCTCACGTCA
CCTGTGAACATCAAGGCAGCAAGACAAGATCTGCTGGCTGTACCCCAAGCTCACGTCA
CCTGTGAACATCAAGGCAGCAAGACAAGATCTGCTGGCTGTACCCCAAGCTCACGTCA

1543 1602
AF055634
UNC5C_exon9
SEQ ID NO:1
SEQ ID NO:3
AAC90914

GCTGCAGCCATGTACAGAGGACCTGTCTATGCCCTGCATGACGTCTCAGACAAAATCCCA
GCTGCAGCCATGTACAGAGGACCTGTCTATGCCCTGCATGACGTCTCAGACAAAATCCCA
GCTGCAGCCATGTACAGAGGACCTGTCTATGCCCTGCATGACGTCTCAGACAAAATCCCA
GCTGCAGCCATGTACAGAGGACCTGTCTATGCCCTGCATGACGTCTCAGACAAAATCCCA

1603 1662
AF055634
UNC5C_exon9
SEQ ID NO:1
SEQ ID NO:3
AAC90914

ATGACCAACTCTCCAATTCTGGATCCACTGCCCAACCTGAAATCAAAGTGTACAACACC
ATGACCAACTCTCCAATTCTGGATCCACTGCCCAACCTGAAATCAAAGTGTACAACACC
ATGACCAACTCTCCAATTCTGGATCCACTGCCCAACCTGAAATCAAAGTGTACAACACC
ATGACCAACTCTCCAATTCTGGATCCACTGCCCAACCTGAAATCAAAGTGTACAACACC
ATGACCAACTCTCCAATTCTGGATCCACTGCCCAACCTGAAATCAAAGTGTACAACACC

1663 1722
AF055634
UNC5C_exon9
SEQ ID NO:1
SEQ ID NO:3
AAC90914

TCAGGTGCTGTCTCCCCCAAGATGACCTCTCTGAGTTTACGTCCAAGCTGTCCCCCTCAG
TCAGGTGCTGTCTCCCCCAAGATGACCTCTCTGAGTTTACGTCCAAGCTGTCCCCCTCAG
TCAGGTGCTGTCTCCCCCAAGATGACCTCTCTGAGTTTACGTCCAAGCTGTCCCCCTCAG
TCAGGTGCTGTCTCCCCCAAGATGACCTCTCTGAGTTTACGTCCAAGCTGTCCCCCTCAG
TCAGGTGCTGTCTCCCCCAAGATGACCTCTCTGAGTTTACGTCCAAGCTGTCCCCCTCAG
TCAAGTGTCTGTCTCCCCCAAGATGACCTCTCTGAGTTTACGTCCAAGCTGTCCCCCTCAG

1723 1782
AF055634

ATGACCCAGTCTGTTGTTGGAGAATGAAGCCCTCAGCCTGAAGAACCAGAGTCTAGCAAGG

- 7 -

UNC5C_exon9
SEQ ID NO:1
SEQ ID NO:3
AAC90914

ATGACCCAGTCGTTGTTGGAGAATGAAGCCCTCAGCCTGAAGAACCAGAGTCTAGCAAGG
ATGACCCAGTCGTTGTTGGAGAATGAAGCCCTCAGCCTGAAGAACCAGAGTCTAGCAAGG
ATGACCCAGTCGTTGTTGGAGAATGAAGCCCTCAGCCTGAAGAACCAGAGTCTAGCAAGG
ATGACCCAGTCGTTGTTGGAGAATGAAGCCCTCAGCCTGAAGAACCAGAGTCTAGCAAGG

AF055634
UNC5C_exon9
SEQ ID NO:1
SEQ ID NO:3
AAC90914

1783 1842
CAGACTGATCCATCCTGTACCGCATTTGGCAGCTTCAACTCGCTGGGAGGTACCTTATT
CAGACTGATCCATCCTGTACCGCATTTGGCAGCTTCAACTCGCTGGGAGGTACCTTATT
CAGACTGATCCATCCTGTACCGCATTTGGCAGCTTCAACTCACTGGGAGGTACCTTATT
CAGACTGATCCATCCTGTACCGCATTTGGCAGCTTCAACTCACTGGGAGGTACCTTATT
CAGACTGATCCATCCTGTACCGCATTTGGCAGCTTCAACTCGCTGGGAGGTACCTTATT

AF055634
UNC5C_exon9
UNC5C_exon10
SEQ ID NO:1
SEQ ID NO:3
AAC90914

1843 1902
GTTCCCAATTTCAGGAGTCAGCTTGCTGATTCCCGCTGGGGCCATTCCCAAGGGAGAGTC
GTTCCCAATTTCAG
GAGTCAGCTTGCTGATTCCCGCTGGGGCCATTCCCAAGGGAGAGTC
GTTCCCAATTTCAGGAGTCAGCTTGCTGATTCCCGCTGGGGCCATTCCCAAGGGAGAGTC
GTTCCCAATTTCAGGAGTCAGCTTGCTGATTCCCGCTGGGGCCATTCCCAAGGGAGAGTC
GTTCCCAATTTCAGGAGTCAGCTTGCTGATTCCCGCTGGGGCCATTCCCAAGGGAGAGTC

AF055634
UNC5C_exon10
UNC5C_exon11
SEQ ID NO:1
SEQ ID NO:3
AAC90914

1903 1962
TACGAAATGTATGTGACTGTACACAGGAAAGAACTATGAGGCCACCCATGGATGACTCT
TACGAAATGTATGTGACTGTACACAGGAAAGAACTATGAG
GCCACCCATGGATGACTCT
TACGAAATGTATGTGACTGTACACAGGAAAGAACTATGAGGCCACCCATGGATGACTCT
TACGAAATGTATGTGACTGTACACAGGAAAGAACTATGAGGCCACCCATGGATGACTCT
TACGAAATGTATGTGACTGTACACAGGAAAGAACTATGAGGCCACCCATGGATGACTCT

AF055634
UNC5C_exon11
SEQ ID NO:1
SEQ ID NO:3
AAC90914

1963 2022
CAGACACTTTTGACCCCTGTGGTGAGCTGTGGGCCCCCAGGAGCTCTGCTCACCCGCCCC
CAGACACTTTTGACCCCTGTGGTGAGCTGTGGGCCCCCAGGAGCTCTGCTCACCCGCCCA
CAGACACTTTTGACCCCTGTGGTGAGCTGTGGGCCCCCAGGAGCTCTGCTCACCCGCCCC
CAGACACTTTTGACCCCTGTGGTGAGCTGTGGGCCCCCAGGAGCTCTGCTCACCCGCCCC
CAGACACTTTTGACCCCTGTGGTGAGCTGTGGGCCCCCAGGAGCTCTGCTCACCCGCCCC

AF055634
UNC5C_exon11
SEQ ID NO:1
SEQ ID NO:3
AAC90914

2023 2082
GTCGTCTCACTATGCATCACTGCGCAGACCCCAATACCGAGGACTGGAAAATACTGCTC
GTCGTCTCACTATGCATCACTGCGCAGACCCCAATACCGAGGACTGGAAAATACTGCTC
GTCGTCTCACTATGCATCACTGCGCAGACCCCAATACCGAGGACTGGAAAATACTGCTC
GTCGTCTCACTATGCATCACTGCGCAGACCCCAATACCGAGGACTGGAAAATACTGCTC
GTCGTCTCACTATGCATCACTGCGCAGACCCCAATACCGAGGACTGGAAAATACTGCTC

AF055634
UNC5C_exon11
UNC5C_exon12
SEQ ID NO:1
SEQ ID NO:3
AAC90914

2083 2142
AAGAACCAGGCAGCACAGGACAGTGGGAGGATGTGGTGGTGGTGGGAGGAAAACTTC
AAGAACCAGGCAGCACAGGACAGTGGGAGGATGTGGTGGTGGTGGGAGGAAAACTTC
GATGTGGTGGTGGTGGGAGGAAAACTTC
AAGAACCAGGCAGCACAGGACAGTGGGAGGATGTGGTGGTGGTGGGAGGAAAACTTC

AF055634
UNC5C_exon12
SEQ ID NO:1
SEQ ID NO:3
AAC90914

2143 2202
ACCAACCCCTGCTACATTAAAGCTGGATGCAGAGGCCTGCCACATCCTCACAGAGAACCTC
ACCAACCCCTGCTACATTAAAGCTGGATGCAGAGGCCTGCCACATCCTCACAGAGAACCTC
ACCAACCCCTGCTACATTAAAGCTGGATGCAGAGGCCTGCCACATCCTCACAGAGAACCTC

AF055634
UNC5C_exon12
SEQ ID NO:1
SEQ ID NO:3
AAC90914

2203 2262
AGCACCTACGCCCTGGTAGGACATTCCACCACCAAGCGGCTGCAAAGCGCCTCAAGCTG
AGCACCTACGCCCTGGTAGGACATTCCACCACCAAGCGGCTGCAAAGCGCCTCAAGCTG
AGCACCTACGCCCTGGTAGGACATTCCACCACCAAGCGGCTGCAAAGCGCCTCAAGCTG

2263 2322

- 8 -

AF055634
UNC5C_exon12
SEQ ID NO:1
SEQ ID NO:3
AAC90914

GCCATCTTTGGGCCCTGTGCTGCTCCTCGCTGGAGTACAGCATCCGAGTCTACTGTCTG
GCCATCTTTGGGCCCTGTGCTGCTCCTCGCTGGAGTACAGCATCCGAGTCTACTGTCTG
GCCATCTTTGGGCCCTGTGCTGCTCCTCGCTGGAGTACAGCATCCGAGTCTACTGTCTG
-----CTCGCTGGAGTACAGCATCCGAGTCTACTGTCTG
-----CTCGCTGGAGTACAGCATCCGAGTCTACTGTCTG

2323 2382
GATGACACCCAGGATGCCCTGAAGGAAATTTACATCTTGAGAGACAGACGGGAGGACAG
GATGACACCCAGGATGCCCTGAAG
GAAATTTTACATCTTGAGAGACAGACGGGAGGACAG
GATGACACCCAGGATGCCCTGAAGGAAATTTACATCTTGAGAGACAGACGGGAGGACAG
GATGACACCCAGGATGCCCTGAAGGAAATTTACATCTTGAGAGACAGACGGGAGGACAG
GATGACACCCAGGATGCCCTGAAGGAAATTTACATCTTGAGAGACAGACGGGAGGACAG

2383 2442
CTCCTAGAAGAACCTAAGGCTCTTCATTTTAAAGGCAGCACCCACAACCTGCGCCTGTCA
CTCCTAGAAGAACCTAAGGCTCTTCATTTTAAAGGCAGCACCCACAACCTGCGCCTGTCA
CTCCTAGAAGAACCTAAGGCTCTTCATTTTAAAGGCAGCACCCACAACCTGCGCCTGTCA
CTCCTAGAAGAACCTAAGGCTCTTCATTTTAAAGGCAGCACCCACAACCTGCGCCTGTCA
CTCCTAGAAGAACCTAAGGCTCTTCATTTTAAAGGCAGCACCCACAACCTGCGCCTGTCA

2443 2502
ATTACGATATCGCCCATTCCTCTGGAAGAGCAAATTGCTGGCTAAATATCAGGAAAT
ATTACGATATCGCCCATTCCTCTGGAAGAGCAAATTGCTGGCTAAATATCAG
GAAAT
ATTACGATATCGCCCATTCCTCTGGAAGAGCAAATTGCTGGCTAAATATCAGGAAAT
ATTACGATATCGCCCATTCCTCTGGAAGAGCAAATTGCTGGCTAAATATCAGGAAAT
ATTACGATATCGCCCATTCCTCTGGAAGAGCAAATTGCTGGCTAAATATCAGGAAAT

2503 2562
CCATTTTACCATGTTTGGAGTGGATCTCAAAGAAACCTGCACCTTCACCTCTGGAA
CCATTTTACCATGTTTGGAGTGGATCTCAAAGAAACCTGCACCTTCACCTCTGGAA
CCATTTTACCATGTTTGGAGTGGATCTCAAAGAAACCTGCACCTTCACCTCTGGAA
CCATTTTACCATGTTTGGAGTGGATCTCAAAGAAACCTGCACCTTCACCTCTGGAA
CCATTTTACCATGTTTGGAGTGGATCTCAAAGAAACCTGCACCTTCACCTCTGGAA

2563 2622
AGATTTAGCCTGAACACAGTGGAGCTGGTTTGCAAACCTCTGTGTGCGGCAGGTGGAAGGA
AGATTTAGCCTGAACACAGTGGAGCTGGTTTGCAAACCTCTGTGTGCGGCAGGTGGAAGGA
AGATTTAGCCTGAACACAGTGGAGCTGGTTTGCAAACCTCTGTGTGCGGCAGGTGGAAGGA
AGATTTAGCCTGAACACAGTGGAGCTGGTTTGCAAACCTCTGTGTGCGGCAGGTGGAAGGA
AGATTTAGCCTGAACACAGTGGAGCTGGTTTGCAAACCTCTGTGTGCGGCAGGTGGAAGGA

2623 2682
GAAGGGCAGATCTTCCAGCTCAACTGCACCGTGTGAGAGGAACCTACTGGCATCGATTG
GAAGGGCAGATCTTCCAGCTCAACTGCACCGTGTGAGAGGAACCTACTGGCATCGATTG
GAACCTACTGGCATCGATTG
GAAGGGCAGATCTTCCAGCTCAACTGCACCGTGTGAGAGGAACCTACTGGCATCGATTG
GAAGGGCAGATCTTCCAGCTCAACTGCACCGTGTGAGAGGAACCTACTGGCATCGATTG
GAAGGGCAGATCTTCCAGCTCAACTGCACCGTGTGAGAGGAACCTACTGGCATCGATTG

2683 2742
CCGCTGCTGGATCCTGCGAACCATCACCACGGTCACGGGGCCAGTGCTTTTACGATC
CCGCTGCTGGATCCTGCGAACCATCACCACGGTCACGGGGCCAGTGCTTTTACGATC
CCGCTGCTGGATCCTGCGAACCATCACCACGGTCACGGGGCCAGTGCTTTTACGATC
CCGCTGCTGGATCCTGCGAACCATCACCACGGTCACGGGGCCAGTGCTTTTACGATC
CCGCTGCTGGATCCTGCGAACCATCACCACGGTCACGGGGCCAGTGCTTTTACGATC

2743 2802
CCTCTCCCTATCCGGCAGAAGCTCTGTAGCAGCCTGGATGCCCCCAGACGAGAGGCCAT
CCTCTCCCTATCCGGCAGAAGCTCTGTAGCAGCCTGGATGCCCCCAGACGAGAGGCCAT
CCTCTCCCTATCCGGCAGAAGCTCTGTAGCAGCCTGGATGCCCCCAGACGAGAGGCCAT
CCTCTCCCTATCCGGCAGAAGCTCTGTAGCAGCCTGGATGCCCCCAGACGAGAGGCCAT
CCTCTCCCTATCCGGCAGAAGCTCTGTAGCAGCCTGGATGCCCCCAGACGAGAGGCCAT

- 9 -

AF055634
UNC5C_exon15
UNC5C_exon16
SEQ ID NO:1
SEQ ID NO:3
AAC90914

2803
GACTGGAGGATGCTGGCCCATAGCTGAACCTGGACAGGTACTTGAATTACTTTGCCACC
GACTGGAGGATGCTGGCCCATAGCTGAACCTGGACAG
GTACTTGAATTACTTTGCCACC
GACTGGAGGATGCTGGCCCATAGCTGAACCTGGACAGGTACTTGAATTACTTTGCCACC
GACTGGAGGATGCTGGCCCATAGCTGAACCTGGACAGGTACTTGAATTACTTTGCCACC
GACTGGAGGATGCTGGCCCATAGCTGAACCTGGACAGGTACTTGAATTACTTTGCCACC

2863
AAATCCAGCCCAACTGGCGTAATCCTGGATCTTTGGGAAGCACAGAACTTCCAGATGGA
AAATCCAGCCCAACTGGCGTAATCCTGGATCTTTGGGAAGCACAGAACTTCCAGATGGA
AAATCCAGCCCAACTGGCGTAATCCTGGATCTTTGGGAAGCACAGAACTTCCAGATGGA
AAATCCAGCCCAACTGGCGTAATCCTGGATCTTTGGGAAGCACAGAACTTCCAGATGGA
AAATCCAGCCCAACTGGCGTAATCCTGGATCTTTGGGAAGCACAGAACTTCCAGATGGA

2923
AACCTGAGCATGCTGGCAGCTGTCTTGAAGAAATGGGAAGACATGAAACGGTGGTGTCC
AACCTGAGCATGCTGGCAGCTGTCTTGAAGAAATGGGAAGACATGAAACGGTGGTGTCC
AACCTGAGCATGCTGGCAGCTGTCTTGAAGAAATGGGAAGACATGAAACGGTGGTGTCC
AACCTGAGCATGCTGGCAGCTGTCTTGAAGAAATGGGAAGACATGAAACGGTGGTGTCC
AACCTGAGCATGCTGGCAGCTGTCTTGAAGAAATGGGAAGACATGAAACGGTGGTGTCC

3028
TTAGCAGCAGAAGGGCAGTATTAAACCACCATGCTGGAAGGGGAAATGCAGTATTAAACCAC
TTAGCAGCAGAAGGGCAGTATTAAACCACCATGCTGGAAGGGGAAATGCAGTATTAAACCAC
TTAGCAGCAGAAGGGCAGTATTAAACCACCATGCTGGAAGGGGAAAT-----
TTAGCAGCAGAAGGGCAGTATTAAACCACCATGCTGGAAGGGGAAAT-----
TTAGCAGCAGAAGGGCAGTATTAAACCACCATGCTGGAAGGGGAAAT-----

3102
CATGCTGGAAGGGGAAATGAAGGACAAAAATGCACAGGGAGTCTGTGGCCGTCACAGTGA
CATGCTGGAAGGGGAAATGAAGGACAAAAATGCACAGGGAGTCTGTGGCCGTCACAGTGA

3162
ATCACAGCTGAGGAGGAAATCCAGATGAGACCAATGCACCTTCACAGGCAAGAAATGCAGCA
ATCACAGCTGAGGAGGAAATCCAGATGAGACCAATGCACCTTCACAGGCAAGAAATGCAGCA

3222
GGAGCCAGAAGGAAAAACAGATACAACCTGCCCATGTACATGCCCACTTTACTCGGAGATCA
GGAGCCAGAAGGAAAAACAGATACAACCTGCCCATGTACATGCCCACTTTACTCGGAGATCA

3282
TCACGGGAGTTAAGAAAAATTTGTAAATTTGTACCTTGAATTTAGCTATCAACCTAATT
TCACGGGAGTTAAGAAAAATTTGTAAATTTGTACCTTGAATTTAGCTATCAACCTAATT

3342
TTCTCTTAGTTGGGCTGTATGCTGTGTGGTACAGGATCTTACAGTTTCTAGGAAACGC
TTCTCTTAGTTGGGCTGTATGCTGTGTGGTACAGGATCTTACAGTTTCTAGGAAACGC

3402
TTTTTATTGCTATCCAGATATATGGATAAACTTTCTTAACAAACCAATTTCTACAAATG
TTTTTATTGCTATCCAGATATATGGATAAACTTTCTTAACAAACCAATTTCTACAAATG

3462
TTGTTTACATCAAATTTGGACAGGGATGCAGACACTGTCCATGGCTCGTTCTATTTTGT
TTGTTTACATCAAATTTGGACAGGGATGCAGACACTGTCCATGGCTCGTTCTATTTTGT

3522
CAATCATTTGAAGTTGAAGCTGTGGACGGTTTGTGTGTCTATTTTCAGATTAGTAATTT

- 10 -

UNC5C_exon16	CAAAATCATTGGAAGTTGAAGCTGTGGACGGTTTGTGTGTCATTTCAGATTAGTAATTT *****
AF055634	3523 3582
UNC5C_exon16	ACAGAGAAATCACAGACTTTTGCTACAAATCGTGTGCATCAAGTGTCTCAGATAATCCTC ACAGAGAAATCACAGACTTTTGCTACAAATCGTGTGCATCAAGTGTCTCAGATAATCCTC *****
AF055634	3583 3642
UNC5C_exon16	CCATCAGTGTCTCTGTTTCTAGAACTTGTAGAACCACTGTACTGTTTGTATCAGGGAAGT CCATCAGTGTCTCTGTTTCTAGAACTTGTAGAACCACTGTACTGTTTGTATCAGGGAAGT *****
AF055634	3643 3702
UNC5C_exon16	GGAGAATCTAAGTGTAAAAAGAAATACTAAGACTCCTATTCTCTGGAGGGACCCCTTCT GGAGAATCTAAGTGTAAAAAGAAATACTAAGACTCCTATTCTCTGGAGGGACCCCTTCT *****
AF055634	3703 3732
UNC5C_exon16	GGTGCCTTTGGGAATAAAGCTGTAGCACTGC GGTGCCTTTGGGAATAAAGCTGTAGCACTGC *****

Table 2 below shows a protein alignment of the IROR α 1-UNC5C (SEQ ID NO: 2) and sROR α 1-UNC5C (SEQ ID NO: 4) fusion proteins with UNC5C:

5 Table 2.

UNC5C exon 1	1 59
UNC5C	-----MRKGLRATAARCGLGLGYLLQMLVLPALALLSASGTGSAAQ
SEQ ID NO:2	-----MRKGLRATAARCGLGLGYLLQMLVLPALALLSASGTGSAAQDDDF
SEQ ID NO:4	MESAPAAPDPAASEPGSSGADAAAGSRETPLNQESARKSEPPAPVRRQSYSTSRDDDF
ROR α 1 exon 1	MESAPAAPDPAASEPGSSGADAAAGSRETPLNQESARKSEPPAPVRRQSYSTSRDDDF MESAPAAPDPAASEPGSSGADAAAGSRETPLNQESARKSEPPAPVRRQSYSTSR *****
UNC5C	60 119
SEQ ID NO:2	FHELPETFPSPDPEPLPHFLIEPEEAYIVKNKPVNLYCKASPATQIYFKCNSEWVHQKDH
SEQ ID NO:4	FHELPETFPSPDPEPLPHFLIEPEEAYIVKNKPVNLYCKASPATQIYFKCNSEWVHQKDH FHELPETFPSPDPEPLPHFLIEPEEAYIVKNKPVNLYCKASPATQIYFKCNSEWVHQKDH *****
UNC5C	120 179
SEQ ID NO:2	IVDERVDETSGLIVREVSIEISRQVEELFGPEDYWCQCVAWSSAGTTKSRKAYVRIAYL
SEQ ID NO:4	IVDERVDETSGLIVREVSIEISRQVEELFGPEDYWCQCVAWSSAGTTKSRKAYVRIAYL IVDERVDETSGLIVREVSIEISRQVEELFGPEDYWCQCVAWSSAGTTKSRKAYVRIAYL *****
UNC5C	180 239
SEQ ID NO:2	RKTFEQEPGLKEVSLEQEVLLQCRPPEGIPVAEVEWLKNEIDIIDPVEDRNFYITIDHNLI
SEQ ID NO:4	RKTFEQEPGLKEVSLEQEVLLQCRPPEGIPVAEVEWLKNEIDIIDPVEDRNFYITIDHNLI RKTFEQEPGLKEVSLEQEVLLQCRPPEGIPVAEVEWLKNEIDIIDPVEDRNFYITIDHNLI *****
UNC5C	240 299
SEQ ID NO:2	IKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGWSTWTEWSVCNSRCGRGYQKRTR
SEQ ID NO:4	IKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGWSTWTEWSVCNSRCGRGYQKRTR IKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGWSTWTEWSVCNSRCGRGYQKRTR *****
UNC5C	300 359
SEQ ID NO:2	TCTNPAPLNGGAFCEGQSVQKIACTTLCFVDGRWTPWSKWSTCGTECTHWRRECTAPAP
SEQ ID NO:4	TCTNPAPLNGGAFCEGQSVQKIACTTLCFVDGRWTPWSKWSTCGTECTHWRRECTAPAP TCTNPAPLNGGAFCEGQSVQKIACTTLCFVDGRWTPWSKWSTCGTECTHWRRECTAPAP *****
UNC5C	360 419
SEQ ID NO:2	KNGGKDCDGLVLQSKNCTDGLCMQ-----TAPDSDDVALYVGIVIA
SEQ ID NO:4	KNGGKDCDGLVLQSKNCTDGLCMQSFITYPISTEQTQNEYGFSSAPDSDDVALYVGIVIA KNGGKDCDGLVLQSKNCTDGLCMQSFITYPISTEQTQNEYGFSSAPDSDDVALYVGIVIA

- 11 -

Fusion specific exon	SFIYPISTEQRTQNEYGFS

UNC5C	420
SEQ ID NO:2	VIVCLAISVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAARQDLLAVPPDLTSAA
SEQ ID NO:4	VIVCLAISVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAARQDLLAVPPDLTSAA

UNC5C	480
SEQ ID NO:2	AMYRGFVYALHDVSDKIPMTNSPILDPLPNLKIKVYNTSGAVSPQDDLSEFTSKLSPQMT
SEQ ID NO:4	AMYRGFVYALHDVSDKIPMTNSPILDPLPNLKIKVYNTSGAVTPQDDLSEFTSKLSPQMT

UNC5C	540
SEQ ID NO:2	QSLLENEALSLKNQSLARQTDPSCTAFGSFNSLGGHLIVPNSGVSLIPAGAIPOGRVYE
SEQ ID NO:4	QSLLENEALSLKNQSLARQTDPSCTAFGSFNSLGGHLIVPNSGVSLIPAGAIPOGRVYE

UNC5C	600
SEQ ID NO:2	MYVTVHRKETMRPFMDSDQTLTPVVS CGPPGALLTRPVVLTMHHCADPNTEDWKILLKN
SEQ ID NO:4	MYVTVHRKETMRPFMDSDQTLTPVVS CGPPGALLTRPVVLTMHHCADPNTEDWKILLKN

UNC5C	660
SEQ ID NO:2	QAAQQWEDVVVVGEENFTTPCYIKLDAEACHILTENLSTYALVGHSTTKAAAKRLKLAI
SEQ ID NO:4	QAAQQWEDVVVVGEENFTTPCYIKLDAEACHILTENLSTYALVGHSTTKAAAKRLKLAI
	GVQHPSLLSG
UNC5C	720
SEQ ID NO:2	FGPLCCSSLEYSIRVYCLDDTQDALKEILHLERQTGGQLLEPKALHFKGSTHNLRLSIH
	FGPLCCSSLEYSIRVYCLDDTQDALKEILHLERQTGGQLLEPKALHFKGSTHNLRLSIH
UNC5C	780
SEQ ID NO:2	DIAHSLWKSLLAKYQEIPFYHVWGSQORNHCTPTLERFSLNTVELVCKLCVRQVEGEG
	DIAHSLWKSLLAKYQEIPFYHVWGSQORNHCTPTLERFSLNTVELVCKLCVRQVEGEG
UNC5C	840
SEQ ID NO:2	QIFQLNCTVSEPTGIDPLLDPAINTITVTGSAFSIPLPIRQKLCSSLDAPQTRGHDW
	QIFQLNCTVSEPTGIDPLLDPAINTITVTGSAFSIPLPIRQKLCSSLDAPQTRGHDW
UNC5C	900
SEQ ID NO:2	RMLAHKLNLDRYLNYPATKSSPTGVILDLEAQNFPDGNLSMLAAVLEEMGRHETVVSLA
	RMLAHKLNLDRYLNYPATKSSPTGVILDLEAQNFPDGNLSMLAAVLEEMGRHETVVSLA
UNC5C	960
SEQ ID NO:2	AEGQY
	AEGQY

As used herein "ROR α 1-UNC5C" refers to both the short and large fusion protein, unless otherwise stated.

The genomic DNA sequence at the fusion site between chromosome 4 (nucleotide positions 1-3221) and chromosome 15 (nucleotide positions 3229-6277) isolated from the translocation 4:15 is set out as SEQ ID NO:11. Nucleotide positions 3222-3228 of SEQ ID NO:11 are a 7bp insertion created in the translocation event. The localisation of this sequence is shown in Fig. 4.

The genomic DNA sequence at the fusion site between chromosome 15 (nucleotide positions 1-998) and chromosome 4 (nucleotide positions 1036-1829) isolated from the

- 12 -

translocation 15:4 is set out as SEQ ID NO:12. Nucleotide positions 999-1035 of SEQ ID NO:12 are a 36bp insertion created in the translocation event. The localisation of this sequence is shown in Fig. 5.

A cDNA encoding a novel isoform of ROR α , designated "ROR α 5", is set out as SEQ ID NO: 7 (with the first nucleotide of the ROR α 5 coding region accorded position 188), and its gene product is set out at SEQ ID NO: 8.

According to one aspect of the present invention there is provided a polypeptide selected from the group consisting of:

- i) a fusion peptide of ROR α 1 – UNC5C polypeptide comprising SEQ ID NO:2 or a fragment thereof of at least 10 amino acids wherein the fragment comprises either the junction of the fusion protein, the flanking regions of the fusion specific exon or the fusion specific exon of SEQ ID NO:2;
- ii) a fusion polypeptide of ROR α 1 – UNC5C polypeptide comprising SEQ ID NO:4 or a fragment thereof of at least 10 amino acids wherein the fragment comprises either the junction of the fusion protein, the flanking regions of the fusion specific exon or the fusion specific exon of SEQ ID:4.
- iii) a polypeptide of ROR α 5 polypeptide comprising the amino acid sequence of SEQ ID NO:8 or a fragment thereof of at least 10 amino acids or an amino acid sequence at least 90%, preferably 95%, 98% or 99% identical to either i) or ii) or iii) or a complement thereof.

In a preferred embodiment, a conservative analogue of the protein is contemplated.

In this specification conservative amino acid analogues of specific amino acid sequences are contemplated which retain the relevant biological properties of a component of the invention but differ in sequence by one or more conservative amino acid substitutions, deletions or additions. However the specifically listed amino acid sequences are preferred. Typical conservative amino acid substitutions are tabulated below.

- 13 -

CONSERVATIVE SUBSTITUTIONS

Original	Exemplary Substitutions	Preferred Substitutions
Ala (A)	Val; Leu; Ile	Val
Arg (R)	Lys; Gln; Asn	Lys
Asn (N)	Gln; His; Lys; Arg	Gln
Asp (D)	Glu	Glu
Cys (C)	Ser	Ser
Gln (Q)	Asn	Asn
Glu (E)	Asp	Asp
Gly (G)	Pro	Pro
His (H)	Asn; Gln; Lys; Arg	Arg
Ile (I)	Leu; Val; Met; Ala; Phe; Norleucine	Leu
Leu (L)	Norleucine; Ile; Val; Met; Ala; Phe	Ile
Lys (K)	Arg; Gln; Asn	Arg
Met (M)	Leu; Phe; Ile	Leu
Phe (F)	Leu; Val; Ile; Ala	Leu
Pro (P)	Gly	Gly
Ser (S)	Thr	Thr
Thr (T)	Ser	Ser
Trp (W)	Tyr	Tyr
Tyr (Y)	Trp; Phe; Thr; Ser	Phe
Val (V)	Ile; Leu; Met; Phe; Ala; Norleucine	Leu

In this specification nucleic acid variations (deletions, substitutions and additions) of
5 specific nucleic acid sequences are contemplated which retain which the ability to hybridise
under stringent conditions to the specific sequence in question. Stringent conditions are
defined as 6xSSC, 0.1% SDS at 60° C for 5 minutes. However specifically listed nucleic acid
sequences are preferred. It is contemplated that peptide nucleic acid may be an acceptable
equivalent of polynucleotide sequences, at least for purposes that do not require translation
10 into protein (Wittung (1994) Nature 368, 561).

- 14 -

By "the junction of the fusion protein", we mean that the fragment of at least 10 amino acids includes the amino acid accorded position 55 and 56 of SEQ ID NO:2 or SEQ ID NO:4.

By "the flanking regions of the fusion specific exon" we mean that the fragment of at least 10 amino acids includes either the amino acids accorded position 383 and 384 or the amino acids accorded position 403 and 404 of SEQ ID NO:2 or SEQ ID NO:4

The invention further features an isolated nucleic acid molecule capable of encoding the ROR α 1-UNC5C polypeptide or a fragment thereof or the ROR α 5 polypeptide or a fragment thereof as defined above or a complement thereof, or a nucleic acid molecule having at least 65% preferably 75%, 85%, 90%, 95%, 98% or 99%, identity thereto. Preferred nucleotide sequences are SEQ ID NO:1, 3 or 7.

The invention further features a vector comprising the nucleic acid molecule as defined above.

Also included within the invention is the vector as defined above, wherein said vector is an expression vector.

The invention further features a host cell comprising the expression vector as defined above.

The invention further features a method of making the ROR α 1-UNC5C polypeptide or the ROR α 5 polypeptide as defined above comprising: incubating the host cell as defined hereinabove in a medium conducive to expression of the polypeptide and optionally purifying the polypeptide.

The invention further features an antibody specific for the ROR α 1-UNC5C polypeptide or the ROR α 5 polypeptide as defined above.

The invention further features a method for detecting the presence of the ROR α 1-UNC5C polypeptide or the ROR α 5 polypeptide as defined above in a biological sample comprising:

contacting the sample with a binding agent capable of specifically binding to the polypeptide, and measuring the amount of binding.

The invention further features a method of detecting the presence of the isolated nucleic acid molecule as defined above comprising:

contacting said nucleic acid with a nucleic acid capable of specifically hybridizing to the translocation breakpoint.

- 15 -

Regions of specificity in the fusion protein/gene include, for example, the junction point and fusion specific insert.

In yet another aspect, the invention features an isolated nucleic acid molecule comprising a nucleotide sequence having at least 65% identity to a degenerate variant of SEQ ID NO: 1, SEQ ID NO: 3 or SEQ ID NO: 7. The invention further features an isolated nucleic acid molecule comprising the complement of SEQ ID NO: 1, SEQ ID NO: 3 or SEQ ID NO: 7.

The invention further features an isolated nucleic acid encoding a ROR α 1-UNC5C polypeptide, the isolated nucleic acid molecule comprising a nucleotide sequence having at least 65% identity to a degenerate variant of SEQ ID NO: 1 or SEQ ID NO: 3.

The invention further features an isolated nucleic acid encoding a ROR α 5 polypeptide, the isolated nucleic acid molecule comprising a nucleotide sequence having at least 65% identity to a degenerate variant of SEQ ID NO: 7.

The invention further features a vector comprising the nucleic acid molecule of any of the above and a host cell comprising the vector.

Also included within the invention is an isolated nucleic acid molecule encoding a protein having 65% sequence identity to the amino acid sequence of SEQ ID NO:2, SEQ ID NO: 4 or SEQ ID NO: 9. Moreover the invention features a purified polypeptide comprising an amino acid sequence having at least 65% identity to the amino acid sequence of SEQ ID NO: 2, SEQ ID NO: 4 or SEQ ID NO: 9.

Also within the invention is a method for producing a protein comprising culturing the host cell having SEQ ID NO:1, SEQ ID NO: 3 or SEQ ID NO:7 under conditions whereby said protein is produced, and recovering the protein from the host cell culture.

The invention further includes a method for detecting a polynucleotide which encodes a protein comprising the amino acid sequence of SEQ ID NO:2, SEQ ID NO: 4 or SEQ ID NO: 8 in a biological sample comprising the steps of:

- a) hybridizing the complement of the polynucleotide sequence which encodes SEQ ID NO: 2, SEQ ID NO: 4 or SEQ ID NO: 8 to a nucleic acid material of a biological sample, thereby forming a hybridization complex; and
- b) detecting the hybridization complex, wherein the presence of the complex correlates with the presence of a polynucleotide encoding a protein in the biological sample.

One aspect of the invention is directed to a method for detecting the presence of other obesity susceptibility genes which may be responsible for the development of obesity by

- 16 -

determining the identity of genes in and around the resulting translocation junction t(4:15) or t(15:4). The method includes determining the identity of genes in and around the translocation junction, e.g., located 20 cM, or less, adjacent to the junction and determining if the gene shows aberrant expression compared to the gene in a person not having the chromosomal translocation.

As used herein, an "obesity susceptibility gene" or "obesity gene" is a gene which as a consequence of the translocation event between chromosome 4 and chromosome 15 has a predisposing influence in the development of obesity in a subject. For example, an obesity susceptibility gene can refer to any or all of the following genes: UNC5C, sUNC5C, ROR α 1, ROR α 5, IROR α 1-UNC5C and sROR α 1-UNC5C.

As used herein "obesity protein" or "obesity susceptibility polypeptide" is the gene product of the obesity susceptibility gene or obesity gene.

In yet another aspect, the invention features a method for detecting the chromosomal translocation comprising analysing a sample of DNA from an individual for the presence of the translocation breakpoint, wherein the presence of the translocation breakpoint indicates that the person has a susceptibility to developing obesity. The nucleic acid sequence of the translocation junctions t(4:15) and t(15:4) is set out in SEQ ID NO: 11 and SEQ ID NO: 12 respectively. The localisation of these sequences are shown in Figs. 4 and 5 respectively.

The invention further includes a method for identifying a test compound that modulates the expression of an obesity susceptibility gene such as UNC5C, sUNC5C, ROR α 1, ROR α 5, IROR α 1-UNC5C and sROR α 1-UNC5C identified in the methods above, comprising contacting a cell capable of expressing an obesity susceptibility gene with a test compound; and determining the level of expression of the obesity susceptibility gene in the presence of the test compound, wherein a decrease or an increase in the obesity susceptibility gene expression, as compared to the level of expression of the obesity susceptibility gene in the absence of the compound, is indicative that the test compound modulates the expression of the obesity susceptibility gene.

Also within the invention is a method of identifying a test compound that modulates the activity of a protein encoded by an obesity susceptibility gene such as UNC5C, sUNC5C, ROR α 1, ROR α 5, IROR α 1-UNC5C and sROR α 1-UNC5C, comprising contacting the protein with a test compound and determining the level of activity of the obesity protein in the presence of the compound, wherein a decrease or an increase in protein activity is indicative that the test compound modulates protein activity.

- 17 -

The invention further features a method of treating a subject having obesity comprising administering an effective amount of the compound identified above. The invention further includes a pharmaceutical composition comprising the compound identified as above, and a pharmaceutically acceptable adjuvant, diluent or carrier.

5 The invention further includes a method of making a pharmaceutical composition. In one aspect, the method includes: contacting a cell capable of expressing an obesity susceptibility gene with a test compound; determining the level of expression of the obesity susceptibility gene identified above in the presence of the test compound, wherein a decrease or an increase in expression of the obesity susceptibility gene, as compared to the level of
10 expression of the obesity susceptibility gene in the absence of the compound, is indicative that the test compound modulates the obesity susceptibility gene expression; and formulating the test compound that modulates the obesity susceptibility gene expression into a pharmaceutical composition.

In another aspect, the invention includes making a pharmaceutical composition
15 including contacting an obesity susceptibility protein with a test compound; determining the level of activity of the obesity susceptibility protein in the presence of the compound, wherein a decrease in the presence of the obesity susceptibility protein, as compared to the level of activity of the obesity susceptibility protein in the absence of the compound, is indicative that the test compound modulates obesity susceptibility protein activity; and formulating the test
20 compound that decreases obesity susceptibility protein activity into a pharmaceutical composition.

In another aspect, the invention includes a method for determining if an obesity susceptibility gene identified above has an altered level of gene expression comprising comparing the level of obesity gene expression in a cell from a patient having obesity with a
25 control cell, and determining the level of expression of the obesity susceptibility gene in both cells, wherein a decrease or an increase in expression of the obesity susceptibility gene, as compared to the level of expression of the obesity susceptibility gene in the control cell, indicates that the obesity susceptibility gene has altered gene expression.

The invention further includes methods of diagnosing obesity or a susceptibility
30 thereto in a subject. The method includes determining the level of an obesity susceptibility gene mRNA and/or the level of an obesity causing protein in a sample from a subject; and comparing the level of the mRNA and/or protein in the sample with a control, wherein a decrease or an increase in the level of the mRNA and/or protein in the sample compared to the

- 18 -

control indicates that the subject has obesity, or a susceptibility thereto. The invention also extends to products useful for carrying out the assay, such as DNA probes (labelled or unlabelled), kits and the like.

The invention further includes methods of diagnosing obesity or a susceptibility thereto in a subject. The method includes analysing for the presence of ROR α 1-UNC5C mRNA and/or protein in a sample from a subject; wherein the presence of the ROR α 1-UNC5C mRNA and/or protein indicates that the subject has obesity, or a susceptibility thereto. The invention also extends to products useful for carrying out the assay, such as DNA probes (labelled or unlabelled), kits and the like.

The present invention also includes gene-based therapies directed at obesity. Therapies may be in the form of polynucleotides comprising all or a portion of the obesity susceptibility gene(s), or obesity gene locus, placed in appropriate vectors or delivered to target cells in direct ways which would modify the function of the obesity protein.

General molecular biology techniques are described in "Current Protocols in Molecular Biology Volumes 1-3, edited by F M Asubel, R Brent and R E Kingston; published by John Wiley, 1998 and Sambrook, J. and Russell, D.W., Molecular Cloning: A Laboratory Manual, the third edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 2001.

Fig. 1 shows a pedigree of the family with schematically drawn chromosomes showing the inheritance of the balanced translocation where all affected individuals are extremely obese. BMI = body mass index.

Fig. 2 shows the normal gene map of human chromosome 4 at the translocation breakpoint.

Fig. 3 shows the normal map of chromosome 15 at the translocation breakpoint with the ROR α locus.

Fig. 4 shows the outcome of the translocation event at chromosome 4 (translocation chromosome 4:15) with the resulting fusion transcripts.

Fig. 5 shows the outcome of the translocation event at chromosome 15 (translocation chromosome 15:4) with the destruction of ROR α 1.

Fig. 6 shows a schematic view of the protein domain architecture of the IROR α 1-UNC5C and sROR α 1-UNC5C fusion protein transcripts together with that of normal human UNC5C.

- 19 -

Fig. 7 shows Taqman[™] analysis of ROR α 5 expression in various tissues, with the highest expression being found in brain. Fig. 7 shows ROR α 5 expression levels where the X axis labelling indicates the following tissues; A = bone marrow, B = brain, C = heart, D = kidney, E = liver, F = lung, G = testis, H = skeletal muscle, I = small intestine, J = spleen and
 5 K = fat.

Fig. 8 shows a schematic view of the 1ROR α 1-UNC5C and sROR α 1-UNC5C fusion transcripts together with UNC5C and sUNC5C.

The invention is based on the finding that a genetic alteration in patients with obesity is the translocation of two specific chromosomes (t(4;15)). This event leads to the production
 10 of two fusion proteins that is thought to promote the development of obesity. It also destroys one copy of ROR α 1 (SEQ ID NOS: 5 & 6) thus maybe reducing the expression of this gene in certain tissues during development. It may also affect the expression of the other isoforms of ROR α , in particular ROR α 5 (SEQ ID NOS: 7 & 8) which is the closest situated isoform of ROR α on the other side of the breakpoint at chromosome 15.

15 The breakpoint in chromosome 4 occurs at cytoband 4q22.3 and the breakpoint in chromosome 15 occurs at cytoband 15q22.2. The balanced translocation results in patients having one normal chromosome 4 and one normal chromosome 15 together with the reciprocal translocation chromosomes. The outcome of the translocation event is that the expression and/or activity of at least one or more genes is affected.

20 The breakpoint at chromosome 4 is between the genes Pyruvate dehydrogenase A2 (PDHA2) and UNC5C (SEQ ID NOS: 9 & 10). While neither of these genes are destroyed in the mutation event, UNC5C was found to be upregulated in adipocytes from obese affected individuals, as compared with control adipocytes.

UNC5 is required for correct projections of pioneer axons and for migrating cells
 25 along the body wall in *C. elegans* (Leung Hagensteijn et. al., Cell 1992 Oct 71 (2): 289-99). UNC5 genetically interacts with UNC6 and UNC40. Mutants of all three genes defects in the axon projections. UNC5 and UNC40 are transmembrane receptors expressed on developing neurons. Both are receptors for the diffusible protein UNC6/Netrin (Hedgecock et. al.; Neuron 1990 Jan ;4(1): 61-85) (Wadsworth; Trends Neurosci 2002 Aug ;25(8): 423-9) Vertebrates
 30 have at least 3 UNC5 like genes named UNC5A, UNC5B and UNC5C (Leonardo et. al.; Nature 1997 Apr ;386 (6627): 833-8) and the protein contains two immunoglobulin domains, two trombospondin domains on the extracellular part and intracellular domains ZU5, DB and

- 20 -

a Death Domain (Leung Hagensteijn et.al., Cell 1992 Oct 71 (2): 289-99) (Keleman & Dickson; Neuron 2001 Nov ;32(4): 605-17). Human UNC5C is localised on chromosome 4 at cytoband q22.3 (Ackerman & Knowles; Genomics 1998 Sep ;52(2): 205-8).

The breakpoint at chromosome 15 is within the first intron of ROR α 1. This isoform of ROR α is destroyed by the mutation event. Moreover, it was found that the translocation resulted in a novel transcript, which is a fusion between ROR α 1 exon 1 with UNC5C exon 2 producing 2 novel fusion proteins through differential splicing within the UNC5C gene (SEQ ID NO:2 and SEQ ID NO: 4).

ROR α 1 is an orphan nuclear hormone receptor localized on chromosome 15 q22.3 (V. Giguere et. al. Genomics 1995 Aug ;28(3): 596-8). Nuclear hormone receptors (NHR) constitute a super family of highly conserved intracellular receptors found in nematodes, flies and vertebrates. The protein architecture is a N-terminal transactivation domain followed by a DNA binding domain, hinge region and a c-terminal ligand binding domain. Many NHRs form dimers upon activation and translocate into the nucleus where they bind DNA to activate or repress target gene expression. There are four publicly available isoforms of ROR α 1, namely ROR α 1-4.

We have identified yet another isoform of ROR α , designated ROR α 5 (SEQ ID 7 & 8). ROR α 5 was isolated with 5'RACE (rapid amplification of C-terminal ends) performed on public available adipocyte cDNA library (clontech). The expression pattern of ROR α 5 is similar to ROR α 1, and is most abundantly expressed in the brain (Fig. 7).

Accordingly, the present invention describes the chromosomal location of obesity susceptibility genes and provides methods for detecting the presence of such genes. The present invention enables the development of novel therapies for obesity by screening for compounds and other entities, such as antibodies, which modulate the activity of the proteins encoded by the associated genes, e.g., UNC5C, sUNC5C, ROR α 1, ROR α 5, lROR α 1-UNC5C and sROR α 1-UNC5C. Knowledge of the sequence of the obesity susceptibility genes also enables the development of novel antigene methods to modulate the expression of the obesity susceptibility gene and also enables the development of novel gene therapy techniques to treat obesity. The discovery of obesity susceptibility genes may also assist in developing novel methods for diagnosing obesity via (i) measuring the levels of the translated mRNA present in affected tissue and (ii) measuring the levels of the protein in affected tissue. It is possible that the diagnosis of obesity, or the susceptibility of an individual to obesity, by these methods

- 21 -

may be achieved in patients who do not yet display the classical symptoms of the disease. Such determination of susceptibility to obesity or the early detection of disease development may lead to earlier clinical intervention, than is currently possible, and may lead to more effective treatment of the disease.

5 *Further identification of obesity susceptibility genes*

The present findings describes a chromosomal translocation which alters the expression/activity of genes in and around the translocation junction causing obesity. The genetic location of such genes are at the breakpoint of a translocation between chromosome 4q22.3 and chromosome 15q22.2. Thus, the present invention includes methods of
10 identifying genes around the translocation breakpoint and determining if any gene around the translocation breakpoint has aberrant expression. The invention includes identifying genes, e.g., 8, 7, 6, 5, 4, 3, 2 or 1 cM around the translocation breakpoint and comparing the expression profile of these genes against the control genes. This can be done by using techniques known in the art to further analyze and delineate this genetic region. For example,
15 using adipocytes isolated from obese individuals carrying the translocation and control adipocytes it is possible to compare the expression profile of an identified obese susceptibility gene identified in the region, or the activity of the gene product of the gene and determine aberrant expression or activity.

Techniques such as Northern analysis, *in situ* hybridization or expression profiling can
20 be used to further verify the association of these genes with obesity. In one example, a reporter-based assay may be devised to detect whether an identified obesity susceptibility gene has a different transcription levels and/or message stability compared to the same gene in a person not susceptible to obesity. Individuals who carry the obesity susceptibility gene may exhibit differences in their ability to regulate protein biosynthesis under different
25 physiological conditions and may display altered physiological abilities.

In another example, the level of obesity gene expression can be assayed by detecting and measuring obesity transcription. For example, RNA from adipocytes from a person carrying the translocation can be isolated and tested utilizing hybridization or PCR techniques. Such analyses can reveal both quantitative and qualitative aspects of the
30 expression pattern of the obesity gene, including activation or inactivation of obesity gene expression.

- 22 -

Pathway mapping

Also within the invention is the identification and elucidation of the obesity biochemical\signal transduction pathway of which the obesity susceptibility protein is a component. In particular, the invention includes identifying the other components of the obesity susceptibility protein biochemical pathway. In this way it is possible to identify the specific critical signaling pathway which links the disease stimulus to the cell's response and enables the identification of new potential targets for therapy intervention.

As used herein, a obesity susceptibility protein "target molecule" is a molecule in the obesity biochemical pathway with which the obesity susceptibility protein binds or interacts, directly or indirectly, with, or is a molecule that regulates the expression of the obesity gene or translation of the protein in nature. For example, the target molecule can be a protein which directly interacts with the obesity susceptibility protein, or can be a protein which does not itself directly interact with the obesity susceptibility protein, but which is a component of the obesity susceptibility protein biochemical pathway. In one example, the target molecule can be an intercellular protein or a protein which facilitates the association of downstream or upstream signaling molecules with the obesity susceptibility protein.

According to a further aspect of the invention there is provided the use of the obesity susceptibility protein in research to identify further gene targets implicated in obesity.

Methods for identifying proteins which interact with the obesity susceptibility protein are known in the art, e.g., the two-hybrid assay (see, e.g., U.S. Pat. No. 5,283,317; Zervos et al. Cell 72:223-232, 1993) or using cell culture techniques to identify binding partners.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for an obesity susceptibility protein is fused to a gene encoding the DNA binding domain of a known transcription factor (e.g., GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, *in vivo*, forming an obesity susceptibility protein - dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (e.g., LacZ) which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell colonies

- 23 -

containing the functional transcription factor can be isolated and used to obtain the cloned gene which encodes the "target" protein which interacts with the obesity susceptibility protein.

Alternatively, binding partners for the obesity susceptibility protein can be identified using cell culture techniques or using cells obtained directly from an obesity patient. The method includes isolating the obesity susceptibility protein of interest from the cell and determining the identity of its target molecule. Initial screening can be accomplished by Western blot analysis to analyse immunochemically, using antibodies against the obesity susceptibility protein, the size of the obesity susceptibility protein-target molecule complex. Further analysis of the complex will reveal the identity of the target molecule.

The gene and protein encoded by the target molecule is also a potential target for therapeutic intervention in obesity disease, for instance in the development of antisense nucleic acid targeted to the mRNA; or more widely in the identification or development of chemical or hormonal therapeutic agents. The person skilled in the art is also capable of devising screening assays to identify compounds (chemical or biological) that modulate (activate or inhibit) the identified gene, which compounds may prove useful as therapeutic agents in treating or preventing obesity.

ROR α 1-UNC5C and ROR α 5 nucleic acid sequence and polypeptide

The invention encompasses a ROR α 1-UNC5C nucleic acid sequence having at least 65%, 70%, 75%, 80%, 85%, 90% or 99%, sequence identity to the ROR α 1-UNC5C nucleic acid sequence of SEQ ID NO: 1 or SEQ ID NO: 3. The invention also encompasses a ROR α 5 nucleic acid sequence having at least 65%, 70%, 75%, 80%, 85%, 90% or 99%, sequence identity to the ROR α 5 nucleic acid sequence of SEQ ID NO: 7. The comparison of sequences and determination of percent sequence identity between two sequences can be accomplished using a mathematical algorithm. In a preferred embodiment, the percent identity between two amino acid sequences is determined using the Needleman and Wunsch (J. Mol. Biol. (48):444-453 (1970)) algorithm which has been incorporated into the GAP program in the GCG software package (available at <http://www.gcg.com>), using either a Blosum 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. In yet another preferred embodiment, the percent identity between two nucleotide sequences is determined using the GAP program in the GCG software package (available at <http://www.gcg.com>), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. In another embodiment, the

- 24 -

percent identity between two amino acid or nucleotide sequences is determined using the algorithm of E. Meyers and W. Miller (CABIOS, 4:11-17 (1989)) which has been incorporated into the ALIGN program (version 2.0), using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4.

5 The invention also encompasses polynucleotides which encode the ROR α 1-UNC5C protein of SEQ ID NO: 2 or SEQ ID NO: 4, or the ROR α 5 protein of SEQ ID NO: 8. It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of nucleotide sequences encoding ROR α 1-UNC5C polypeptide or ROR α 5 polypeptide, some bearing minimal homology to the nucleotide sequences of any known and
10 naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of nucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the nucleotide sequence of naturally occurring ROR α 1-UNC5C polypeptide or ROR α 5 polypeptide, and all such variations are to be considered as
15 being specifically disclosed.

The invention also encompasses production of DNA sequences, or fragments thereof, which encode ROR α 1-UNC5C, sROR α 1-UNC5C or ROR α 5 polypeptide and their derivatives, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents that
20 are well known in the art.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed ROR α 1-UNC5C or ROR α 5 nucleic acid sequence, and in particular, those shown in SEQ ID NO: 1, SEQ ID NO: 3 or SEQ ID NO: 7, under various conditions of stringency as taught in Wahl, G. M. and S. L. Berger (1987; Methods Enzymol. 152:399-407) and Kimmel, A. R. (1987; Methods Enzymol. 152:507-511). Appropriate
25 stringency conditions which promote DNA hybridization, for example, 6.0 X sodium chloride/sodium citrate (SSC) at about 45°C, followed by a wash of 2.0 X SSC at 50°C, are known to those skilled in the art or can be found in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. For example, the salt concentration in the wash
30 step can be selected from a low stringency of about 2.0 SSC at 50°C to a high stringency of about 0.2 X SSC at 50°C. In addition, the temperature in the wash step can be increased from low stringency conditions at room temperature, about 22° C, to high stringency conditions at

- 25 -

about 65°C. Moderately stringent conditions are, for example at about 2.0 X SSC and about 40°C.

Also included in the invention are ROR α 1-UNC5C polypeptides having at least 65% amino acid sequence identity to the ROR α 1-UNC5C polypeptide of SEQ ID NO: 2 or SEQ ID NO: 4 and which retains at least one biological or other functional characteristic or activity of the ROR α 1-UNC5C polypeptide. A most preferred ROR α 1-UNC5C variant is one having at least 65% 70%, 80% or 90% amino acid sequence identity to SEQ ID NO: 2 or SEQ ID NO: 4.

Also included in the invention is ROR α 5 polypeptide having at least 65% amino acid sequence identity to the ROR α 5 polypeptide of SEQ ID NO: 8 and which retains at least one biological or other functional characteristic or activity of the ROR α 5 polypeptide. A most preferred ROR α 5 variant is one having at least 65% 70%, 80% or 90% amino acid sequence identity to SEQ ID NO: 8.

The invention also includes variants of the ROR α 1-UNC5C polypeptide or the ROR α 5 polypeptide which can contain one or more substitutions of amino acid residues which result in a silent change and a functionally equivalent ROR α 1-UNC5C polypeptide or ROR α 5 polypeptide respectively. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues as long as the biological or immunological activity of ROR α 1-UNC5C polypeptide or ROR α 5 polypeptide is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid; positively charged amino acids may include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values may include leucine, isoleucine, and valine, glycine and alanine, asparagine and glutamine, serine and threonine, and phenylalanine and tyrosine.

Obesity Proteins and Polypeptides

Obesity proteins such as the ROR α 1, ROR α 5, UNC5C, sUNC5C or the ROR α 1-UNC5C fusion proteins can be prepared for a variety of uses, including but not limited to the generation of antibodies, as reagents in diagnostic assays, in the identification of other cellular gene products involved in the pathogenesis of obesity, as reagents in assays for screening for compounds that can be used in the treatment of obesity, and as pharmaceutical reagents useful in the treatment of obesity.

A variety of host-expression vector systems may be utilized to express the obesity

- 26 -

nucleotide sequences of the invention. The expression systems that may be used for purposes of the invention include but are not limited to microorganisms such as bacteria (e.g., *E. coli*, *B. subtilis*) transformed with recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vectors containing obesity nucleotide sequences or mammalian cell systems (e.g., COS, CHO, BHK, 293, 3T3) harbouring recombinant expression constructs containing promoters derived from the genome of mammalian cells (e.g., metallothionein promoter) or from mammalian viruses (e.g., the adenovirus late promoter; the vaccinia virus 7.5K promoter).

In bacterial systems, a number of expression vectors may be advantageously selected depending upon the use intended for the obesity gene product being expressed. For example, when a large quantity of such a protein is to be produced, for the generation of pharmaceutical compositions of obesity protein or for raising antibodies to the obesity protein, for example, vectors which direct the expression of high levels of fusion protein products that are readily purified may be desirable. Such vectors include, but are not limited to, the *E. coli* expression vector pUR278 (Ruther *et al.*, EMBO J. 2:1791, 1983), in which the obesity coding sequence may be ligated individually into the vector in frame with the lacZ coding region so that a fusion protein is produced; pIN vectors (Inouye & Inouye, Nucleic Acids Res. 13:3101-3109, 1985; Van Heeke & Schuster, J. Biol. Chem. 264:5503-5509, 1989); and the like. PGEX vectors may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. The PGEX vectors are designed to include thrombin or factor Xa protease cleavage sites so that the cloned target gene product can be released from the GST moiety.

In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, the obesity nucleotide sequence of interest may be ligated to an adenovirus transcription/translation control complex, e.g., the late promoter and tripartite leader sequence. This chimeric gene may then be inserted in the adenovirus genome by *in vitro* or *in vivo* recombination. Insertion in a non-essential region of the viral genome (e.g., region E1 or E3) will result in a recombinant virus that is viable and capable of expressing the obesity gene product in infected hosts. (E.g., See Logan & Shenk, Proc. Natl. Acad. Sci. USA 81:3655-3659, 1984).

For long-term, high-yield production of recombinant proteins, stable expression is

- 27 -

preferred. For example, cell lines which stably express the obesity sequences described above may be engineered. Rather than using expression vectors which contain viral origins of replication, host cells can be transformed with DNA controlled by appropriate expression control elements (e.g., promoter, enhancer sequences, transcription terminators, polyadenylation sites, etc.), and a selectable marker. Following the introduction of the foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched media, and then are switched to a selective media. The selectable marker in the recombinant plasmid confers resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci which in turn can be cloned and expanded into cell lines.

10 This method may advantageously be used to engineer cell lines which express the obesity gene product. Such engineered cell lines may be particularly useful in screening and evaluation of compounds that affect the endogenous activity of the obesity gene product.

Alternatively, any fusion protein may be readily purified by utilizing an antibody specific for the fusion protein being expressed. For example, a system described by Janknecht *et al.* allows for the ready purification of non-denatured fusion proteins expressed in human cell lines (Janknecht, *et al.*, Proc. Natl. Acad. Sci. USA 88:8972-8976, 1991). In this system, the gene of interest is subcloned into a vaccinia recombination plasmid such that the gene's open reading frame is translationally fused to an amino-terminal tag consisting of six histidine residues. Extracts from cells infected with recombinant vaccinia virus are loaded onto Ni.²⁺-nitriloacetic acid-agarose columns and histidine-tagged proteins are selectively eluted with imidazole-containing buffers.

20

The obesity gene products can also be expressed in transgenic animals. Animals of any species, including, but not limited to, mice, rats, rabbits, guinea pigs, pigs, micro-pigs, goats, and non-human primates, for example, baboons, monkeys, and chimpanzees may be used to generate obesity transgenic animals.

25

Any technique known in the art may be used to introduce the obesity transgene into animals to produce the founder lines of transgenic animals. Such techniques include, but are not limited to pronuclear microinjection (U.S. Pat. No. 4,873,191); retrovirus mediated gene transfer into germ lines (Van der Putten *et al.*, Proc. Natl. Acad. Sci. USA 82:6148-6152, 1985); gene targeting in embryonic stem cells (Thompson *et al.*, Cell 56:313-321, 1989); electroporation of embryos (Lo, Mol. Cell. Biol. 3:1803-1814, 1983); and sperm-mediated gene transfer (Lavitrano *et al.*, Cell 57:717-723, 1989); etc. For a review of such techniques, see Gordon, Transgenic Animals, Intl. Rev. Cytol. 115:171-229, 1989, which is incorporated

30

- 28 -

by reference herein in its entirety.

The present invention provides for transgenic animals that carry the obesity transgene in all their cells, as well as animals which carry the transgene in some, but not all their cells, i.e., mosaic animals. The transgene may be integrated as a single transgene or in concatamers, e.g., head-to-head tandems or head-to-tail tandems. The transgene may also be selectively introduced into and activated in a particular cell type by following, for example, the teaching of Lasko *et al.* (Lasko, M. *et al.*, Proc. Natl. Acad. Sci. USA 89:6232-6236, 1992). Once transgenic animals have been generated, the expression of the recombinant obesity gene may be assayed utilizing standard techniques. Initial screening may be accomplished by Southern blot analysis or PCR techniques to analyze animal tissues to assay whether integration of the transgene has taken place. The level of mRNA expression of the transgene in the tissues of the transgenic animals may also be assessed using techniques which include but are not limited to Northern blot analysis of tissue samples obtained from the animal, in situ hybridization analysis, and RT-PCR. Samples of obesity gene-expressing tissue, may also be evaluated immunocytochemically using antibodies specific for the obesity transgene product.

Antibodies to obesity Proteins

Antibodies that specifically recognize one or more epitopes of obesity proteins, or peptide fragments, are also encompassed by the invention. Such antibodies include but are not limited to polyclonal antibodies, monoclonal antibodies (mAbs), humanized or chimeric antibodies, single chain antibodies, Fab fragments, F(ab')₂ fragments, fragments produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies, and epitope-binding fragments of any of the above.

The antibodies of the invention may be used, for example, in the detection of the obesity in a biological sample and may, therefore, be utilized as part of a diagnostic or prognostic technique whereby patients may be tested for abnormal amounts of obesity. Such antibodies may also be utilized in conjunction with, for example, compound screening schemes for the evaluation of the effect of test compounds on expression and/or activity of the obesity gene product. Additionally, such antibodies can be used in conjunction with the gene therapy techniques described, for example, evaluate the normal and/or engineered obesity - expressing cells prior to their introduction into the patient. Such antibodies may additionally be used as a method for the inhibition of abnormal obesity activity. Thus, such antibodies may, therefore, be utilized as part of obesity treatment methods.

Methods of making and detecting labelled antibodies are well known (Campbell;

- 29 -

Monoclonal Antibody Technology, in: Laboratory Techniques in Biochemistry and Molecular Biology, Volume 13. Eds: Burdon R *et al.* Elsevier, Amsterdam (1984)). The term antibody includes both monoclonal antibodies, which are a substantially homogeneous population, and polyclonal antibodies which are heterogeneous populations. The term also
5 includes inter alia, humanised and chimeric antibodies. Monoclonal antibodies to specific antigens may be obtained by methods known to those skilled in the art, such as from hybridoma cells, phage display libraries or other methods. Monoclonal antibodies may be inter alia, human, rat or mouse derived. For the production of human monoclonal antibodies, hybridoma cells may be prepared by fusing spleen cells from an immunised animal, e.g. a
10 mouse, with a tumour cell. Appropriately secreting hybridoma cells may thereafter be selected (Koehler & Milstein, Nature 256:495-497 (1975); Cole *et al.*, "Monoclonal antibodies and Cancer Therapy", Alan R Liss Inc, New York N.Y. pp 77-96 (1985)). Such antibodies may be of any immunoglobulin class including IgG, IgM, IgE, IgA, IgD and any subclass thereof.

15 Polyclonal antibodies can be generated by immunisation of an animal (such as a mouse, rat, goat, horse, sheep etc) with an antigen, such as one of the obesity proteins used in this invention.

Rodent antibodies may be humanised using recombinant DNA technology according to techniques known in the art. Alternatively, chimeric antibodies, single chain antibodies,
20 Fab fragments may also be developed against the polypeptides of the invention (Huse *et al.*, Science 256:1275-1281 (1989)), using skills known in the art. Antibodies so produced have a number of uses which will be evident to the molecular biologist or immunologist skilled in the art. Such uses include, but are not limited to, monitoring enzyme expression, development of assays to measure enzyme activity and use as a therapeutic agent. Enzyme linked
25 immunosorbant assays (ELISAs) are well known in the art and would be particularly suitable for detecting the obesity protein or polypeptide fragments thereof in a test sample.

Screening Assay

The invention also provides a method for identifying modulators, i.e., test compounds
30 (e.g., peptides, peptidomimetics, small molecules or other drugs) which modulate the activity of an obesity susceptibility gene.

In one example, the invention provides methods for screening for test compounds for use in the treatment of obesity by screening for test compounds that modulate the activity of

- 30 -

the obesity protein, or a portion thereof. The test compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries. The method can be a cell-based method or a cell free method. The screening methods according to the invention may be operated using
5 conventional procedures, for example by bringing the test compound or compounds to be screened and an appropriate substrate into contact with the obesity polypeptide, or a cell capable of producing it, or a cell membrane preparation thereof, and determining affinity for the obesity polypeptide in accordance with standard techniques.

Any compound identified in this way may prove useful in the treatment of obesity in
10 humans and/or other animals. The invention thus extends to a compound selected through its ability to regulate the activity of the obesity protein *in vivo* as primarily determined in a screening assay utilising an obesity polypeptide or a homologue or fragment thereof, or a gene coding therefore for use in the treatment of a disease in which the over- or under-activity or unregulated activity of the protein is implicated.

15 According to a further aspect of the invention there is provided a screening assay or method for identifying potential anti- obesity therapeutic compounds comprising contacting an assay system capable of detecting the effect of a test compound against expression level of obesity, with a test compound and assessing the change in expression level of obesity. Compounds that modulate the expression of DNA or RNA of obesity polypeptides may be
20 detected by a variety of assay systems. A suitable assay system may be a simple "yes/no" assay to determine whether there is a change in expression of a reporter gene, such as beta-galactosidase, luciferase, green fluorescent protein or others known to the person skilled in the art (reviewed by Naylor, Biochem. Pharmacol. 58:749-57,1999). The assay system may be made quantitative by comparing the expression or function of a test sample with the levels of
25 expression or function in a standard sample. Systems in which transcription factors are used to stimulate a positive output, such as transcription of a reporter gene, are generally referred to as "one-hybrid systems" (Wang, M.M. and Reed, R.R. Nature 364:121-126, 1993). Using a transcription factor to stimulate a negative output (growth inhibition) may thus be referred to as a "reverse one-hybrid system" (Vidal et al, *supra*). Therefore, in an embodiment of the
30 present invention, a reporter gene is placed under the control of the obesity promoter.

In a further aspect of the invention we provide a cell or cell line comprising a reporter gene under the control of the obesity promoter.

- 31 -

According to another aspect of the present invention there is provided a method of screening for a compound potentially useful for treatment of obesity which comprises assaying the compound for its ability to modulate the activity or amount of obesity.

Preferably the assay is selected from:

- 5 i) measurement of obesity activity using a cell line which expresses the obesity polypeptide or using purified obesity polypeptide; and
- ii) measurement of obesity transcription or translation in a cell line expressing the obesity polypeptide.

Thus, in a further aspect of the invention, cell cultures expressing the obesity polypeptide
10 can be used in a screen for therapeutic agents. Effects of test compounds may be assayed by changes in mRNA or protein of obesity. As described above, cells (i.e. mammalian, bacterial etc) can be engineered to express the obesity polypeptide.

Thus, according to a further aspect of the invention there is provided a method of testing potential therapeutic agents for the ability to suppress the obesity phenotype comprising
15 contacting a test compound with a cell engineered to express the obesity polypeptide; and determining whether said test compound suppressed expression of the obesity polypeptide.

We also provide a method for identifying inhibitors of transcription of obesity, which method comprises contacting a potential therapeutic agent with a cell or cell line as described above and determining inhibition of obesity transcription by the potential therapeutic agent by
20 reference to a lack of or reduction in expression of the reporter gene.

Any convenient test compound or library of test compounds may be used in conjunction with the test assay. Particular test compounds include low molecular weight chemical compounds (preferably with a molecular weight less than 1500 daltons) suitable as pharmaceutical or veterinary agents for human or animal use, or compounds for non-
25 administered use such as cleaning/sterilising agents or for agricultural use. Test compounds may also be biological in nature, such as antibodies.

According to a further aspect of the invention there is provided a compound identified by a screening method as defined herein.

According to another aspect of the present invention there is provided use of a compound
30 able to modulate the activity or amount of obesity in the preparation of a medicament for the treatment of obesity. Modulation of the amount of obesity by a compound may be brought about for example through altered gene expression level or message stability. Modulation of the activity of obesity by a compound may also be brought about for example through

- 32 -

compound binding to the obesity protein. In one embodiment, modulation of obesity comprises use of a compound able to reduce the activity or amount of obesity. In another embodiment, modulation of obesity comprises use of a compound able to increase the activity or amount of obesity.

5 *Diagnostic test*

To determine if an individual is susceptible to obesity determination of the presence of the translocation junction between chromosome 4q22.3 and chromosome 15q22.2 can be performed to predict whether the individual has a high risk obesity chromosome. Moreover, determination of the translocation breakpoint can be used to identify the underlying cause of
10 obesity in a person who is obese.

A diagnostic test method lends itself readily to the formulation of kits which can be utilized in diagnosis. The diagnostic kits may comprise appropriate packaging and instructions for use in the methods of the invention. Such kits may further comprise appropriate buffer(s) and polymerase(s) such as thermostable polymerases, for example taq
15 polymerase.

Methods for the diagnosis of obesity may, for example, utilize reagents such as the nucleotide sequences which span the breakpoint of sROR α 1 -UNC5C or lROR α 1 -UNC5C antibodies. Specifically, such reagents may be used, for example, for: (1) the detection of the presence of short or long ROR α 1 -UNC5C transcript, (2) the detection of perturbations or
20 abnormalities in the signal transduction pathway mediated by ROR α 1 -UNC5C or UNC5C.

For the detection of mutations, any nucleated cell can be used as a starting source for genomic nucleic acid. For the detection of an obesity susceptibility gene such as sROR α 1 -UNC5C or lROR α 1-UNC5C gene products, any cell type or tissue in which the obesity gene is expressed can be used.

25 *Detection of the obesity Gene and Transcripts*

Obesity susceptibility gene transcripts can be detected by utilizing a number of techniques. Nucleic acid from any nucleated cell can be used as the starting point for such assay techniques, and may be isolated according to standard nucleic acid preparation procedures which are well known to those of skill in the art.

30 DNA may be used in hybridization or amplification assays of biological samples to detect abnormalities involving obesity gene structure or gene regulatory elements, including point mutations, insertions, deletions and chromosomal rearrangements. Such assays may include, but are not limited to, Southern analyses, single stranded conformational polymorphism

- 33 -

analyses (SSCP), and PCR analyses.

Such diagnostic methods for the detection of obesity gene-specific mutations can involve for example, contacting and incubating nucleic acids including recombinant DNA molecules, cloned genes or degenerate variants thereof, obtained from a sample, for example, derived
5 from a patient sample or other appropriate cellular source, with one or more labelled nucleic acid reagents including recombinant DNA molecules, under conditions favourable for the specific annealing of these reagents to their complementary sequences within the obesity gene. Preferably, the lengths of these nucleic acid reagents are at least 15 to 30 nucleotides. After incubation, all non-annealed nucleic acids are removed from the nucleic acid: obesity
10 molecule hybrid. The presence of nucleic acids which have hybridized, if any such molecules exist, is then detected. Using such a detection scheme, the nucleic acid from the cell type or tissue of interest can be immobilized, for example, to a solid support such as a membrane, or a plastic surface such as that on a microtiter plate or polystyrene beads. Detection of the remaining, annealed, labelled obesity nucleic acid reagents is accomplished using standard
15 techniques well-known to those in the art. The obesity gene sequences to which the nucleic acid reagents have annealed can be compared to the annealing pattern expected from a normal obesity gene sequence in order to determine whether an obesity gene mutation is present.

Detection of the obesity Gene Products

Antibodies directed against wild type or mutant obesity gene products or conserved
20 variants or peptide fragments thereof, can also be used as obesity diagnostics. Such diagnostic methods, can be used to detect abnormalities in the level of obesity gene expression, or abnormalities in the structure and/or temporal, tissue, cellular, or subcellular location of the obesity, and may be performed *in vivo* or *in vitro*, such as, for example, on biopsy tissue.

25 For example, antibodies directed to epitopes of the obesity can be used *in vivo* to detect the pattern and level of expression of the obesity in the body. Such antibodies can be labelled, and injected into a subject in order to visualize binding to the obesity expressed in the body using methods such as X-rays, CAT-scans, or MRI.

Additionally, any obesity fusion protein or obesity conjugated protein whose presence
30 can be detected, can be administered. For example, obesity fusion or conjugated proteins labelled with a radio-opaque or other appropriate compound can be administered and visualized *in vivo*, as discussed, above for labelled antibodies.

- 34 -

Alternatively, immunoassays or fusion protein detection assays, as described above, can be utilized on biopsy and autopsy samples *in vitro* to permit assessment of the expression pattern of the obesity. Such assays are not confined to the use of antibodies that define the obesity, but can include the use of antibodies directed to epitopes of any region of the obesity.

5 The tissue or cell type to be analyzed will generally include those which are known, or suspected, to express the obesity gene. The protein isolation methods employed herein can, for example, be such as those described in Harlow and Lane (Harlow, E. and Lane, D., 1988, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.), which is incorporated herein by reference in its entirety. The isolated cells can
10 be derived from cell culture or from a patient. The analysis of cells taken from culture may be a necessary step in the assessment of cells that could be used as part of a cell-based gene therapy technique or, alternatively, to test the effect of compounds on the expression of the obesity gene.

Gene therapy

15 The invention encompasses gene therapy methods and compositions for treating and preventing obesity.

In one example, the loss of normal obesity gene product function can result in the development of obesity. Increasing obesity gene product activity, or activation of the obesity pathway (e.g., downstream activation) would therefore facilitate progress in individuals
20 exhibiting a deficient level of obesity gene expression and/or obesity activity.

Alternatively, obesity may be ameliorated by decreasing the level of obesity gene expression, and/or obesity gene activity, and/or downregulating activity of the obesity pathway (e.g., by targeting downstream signalling events). Different approaches are discussed below.

25 *Inhibition of obesity Expression or obesity Activity*

Any method that neutralizes or inhibits expression of the obesity gene (either transcription or translation), e.g. ROR α 1, ROR α 5 and ROR α 1-UNC5C, can be used to prevent or treat obesity.

For example, the administration of soluble peptides, proteins, fusion proteins, or
30 antibodies (including anti-idiotypic antibodies) that bind to and "neutralize" circulating obesity can be used. Such obesity neutralizing peptides, proteins, fusion proteins, anti-idiotypic antibodies or Fabs are administered to a subject in amounts sufficient to treat or prevent obesity.

- 35 -

In an alternate embodiment, therapy can be designed to reduce the level of endogenous obesity gene expression, for example, using antisense or ribozyme approaches to inhibit or prevent translation of obesity mRNA-transcripts; triple helix approaches to inhibit transcription of the obesity gene; or targeted homologous recombination to inactivate or "knock out" the obesity gene or its endogenous promoter. Antisense, ribozyme or DNA constructs can be administered directly to the site containing the target cells; or can be directed to the target cells.

Antisense approaches involve the design of oligonucleotides (either DNA or RNA) that are complementary to obesity mRNA. The antisense oligonucleotides will bind to the complementary obesity mRNA transcripts and prevent translation. Absolute complementarity, although preferred, is not required. A sequence "complementary" to a portion of an RNA, as referred to herein, means a sequence having sufficient complementarity to be able to hybridize with the RNA, forming a stable duplex; in the case of double-stranded antisense nucleic acids, a single strand of the duplex DNA may thus be tested, or triplex formation may be assayed. The ability to hybridize will depend on both the degree of complementarity and the length of the antisense nucleic acid. Generally, the longer the hybridizing nucleic acid, the more base mismatches with an RNA it may contain and still form a stable duplex (or triplex, as the case may be). One skilled in the art can ascertain a tolerable degree of mismatch by use of standard procedures to determine the melting point of the hybridized complex.

Oligonucleotides that are complementary to the 5' end of the message, for example, the 5' untranslated sequence up to and including the AUG initiation codon, should work most efficiently at inhibiting translation. However, sequences complementary to the 3' untranslated sequences of mRNAs have been shown to be effective at inhibiting translation of mRNAs as well. Thus, oligonucleotides complementary to either the 5'- or 3'-non-translated, non-coding regions of the obesity can be used in an antisense approach to inhibit translation of endogenous obesity mRNA. Oligonucleotides complementary to the 5' untranslated region of the mRNA should include the complement of the AUG start codon. Antisense oligonucleotides complementary to mRNA coding regions are less efficient inhibitors of translation but could be used in accordance with the invention. Whether designed to hybridize to the 5'-, 3'- or coding region of obesity mRNA, antisense nucleic acids should be at least six nucleotides in length, and are preferably oligonucleotides ranging from 6 to about 50 nucleotides in length. In specific aspects, the oligonucleotide is at least 10 nucleotides, at least 17 nucleotides, at least 25 nucleotides, or at least 50 nucleotides in length.

- 36 -

Regardless of the choice of target sequence, it is preferred that *in vitro* studies are first performed to quantitate the ability of the antisense oligonucleotide to inhibit gene expression. It is preferred that these studies utilize controls that distinguish between antisense gene inhibition and nonspecific biological effects of oligonucleotides. It is also preferred that these studies compare levels of the target RNA or protein with that of an internal control RNA or protein. Additionally, it is envisioned that results obtained using the antisense oligonucleotide are compared with those obtained using a control oligonucleotide. It is preferred that the control oligonucleotide is of approximately the same length as the test oligonucleotide and that the nucleotide sequence of the oligonucleotide differs from the antisense sequence no more than is necessary to prevent specific hybridization to the target sequence.

While antisense nucleotides complementary to the obesity coding region sequence could be used, those complementary to the transcribed untranslated region are most preferred. The antisense molecules should be delivered to cells which express the obesity *in vivo*. A number of methods have been developed for delivering antisense DNA or RNA to cells; for example, antisense molecules can be injected directly into the tissue site, or modified antisense molecules, designed to target the desired cells (e.g., antisense linked to peptides or antibodies that specifically bind receptors or antigens expressed on the target cell surface) can be administered systemically.

Ribozyme molecules designed to catalytically cleave obesity mRNA transcripts can also be used to prevent translation of obesity mRNA and expression of obesity. (See, e.g., PCT International Publication WO90/11364, published Oct. 4, 1990; Sarver *et al.*, 1990, Science 247:1222-1225). While ribozymes that cleave mRNA at site specific recognition sequences can be used to destroy obesity mRNAs, the use of hammerhead ribozymes is preferred. Hammerhead ribozymes cleave mRNAs at locations dictated by flanking regions that form complementary base pairs with the target mRNA. The sole requirement is that the target mRNA have the following sequence of two bases: 5'-UG-3'. The construction and production of hammerhead ribozymes is well known in the art and is described more fully in Haseloff and Gerlach, 1988, Nature 334:585-591. There are hundreds of potential hammerhead ribozyme cleavage sites within the nucleotide sequence of human obesity cDNA. Preferably the ribozyme is engineered so that the cleavage recognition site is located near the 5' end of the OBESITY mRNA; i.e., to increase efficiency and minimize the intracellular accumulation of non-functional mRNA transcripts.

- 37 -

The ribozymes of the present invention also include RNA endoribonucleases (hereinafter "Cech-type ribozymes") such as the one which occurs naturally in *Tetrahymena Thermophila* (known as the IVS, or L-19 IVS RNA) and which has been extensively described by Thomas Cech and collaborators (Zaug, *et al.*, 1984, *Science* 224:574-578; Zaug and Cech, 1986, *Science* 231:470-475; Zaug, *et al.*, 1986, *Nature* 324:429-433; published International patent application No. WO 88/04300 by University Patents Inc.; Been and Cech, 1986, *Cell* 47:207-216). The Cech-type ribozymes have an eight basepair active site that hybridizes to a target RNA sequence, whereafter cleavage of the target RNA takes place. The invention encompasses those Cech-type ribozymes that target eight basepair active site sequences that are present in obesity.

As in the antisense approach, the ribozymes can be composed of modified oligonucleotides (e.g. for improved stability, targeting, etc.) and should be delivered to cells which express the obesity *in vivo*. A preferred method of delivery involves using a DNA construct "encoding" the ribozyme under the control of a strong constitutive pol III or pol II promoter, so that transfected cells will produce sufficient quantities of the ribozyme to destroy endogenous obesity messages and inhibit translation. Because ribozymes, unlike antisense molecules, are catalytic, a lower intracellular concentration is required for efficiency.

In an alternative embodiment for neutralizing circulating obesity, cells that are genetically engineered to express such soluble or secreted forms of obesity can be administered to a patient, whereupon they will serve as "bioreactors" *in vivo* to provide a continuous supply of the neutralizing protein.

Endogenous obesity gene expression can also be reduced by inactivating or "knocking out" the obesity gene or its promoter using targeted homologous recombination. (E.g., see Smithies *et al.*, 1985, *Nature* 317:230-234; Thomas & Capecchi, 1987, *Cell* 51:503-512; Thompson *et al.*, 1989, *Cell* 5:313-321; each of which is incorporated by reference herein in its entirety).

In yet another embodiment of the invention, the activity of obesity can be reduced using a "dominant negative" approach to effectuate obesity. To this end, constructs that encode defective obesity can be used in gene therapy approaches to diminish the activity of the obesity in appropriate target cells.

Pharmaceutical composition containing the Identified Candidate Compound

The compounds having the desired activity may be administered in a physiologically acceptable carrier to an obesity patient. Such compositions of the invention may be in a form

- 38 -

suitable for oral use (for example as tablets, lozenges, hard or soft capsules, aqueous or oily suspensions, emulsions, dispersible powders or granules, syrups or elixirs), for topical use (for example as creams, ointments, gels, or aqueous or oily solutions or suspensions), for administration by inhalation (for example as a finely divided powder or a liquid aerosol), for administration by insufflation (for example as a finely divided powder) or for parenteral administration (for example as a sterile aqueous or oily solution for intravenous, subcutaneous, intramuscular or intramuscular dosing or as a suppository for rectal dosing).

The compositions of the invention may be obtained by conventional procedures using conventional pharmaceutical excipients, well known in the art. Thus, compositions intended for oral use may contain, for example, one or more colouring, sweetening, flavouring and/or preservative agents.

Suitable pharmaceutically acceptable excipients for a tablet formulation include, for example, inert diluents such as lactose, sodium carbonate, calcium phosphate or calcium carbonate, granulating and disintegrating agents such as corn starch or algenic acid; binding agents such as starch; lubricating agents such as magnesium stearate, stearic acid or talc; preservative agents such as ethyl or propyl p-hydroxybenzoate, and anti-oxidants, such as ascorbic acid. Tablet formulations may be uncoated or coated either to modify their disintegration and the subsequent absorption of the active ingredient within the gastrointestinal track, or to improve their stability and/or appearance, in either case, using conventional coating agents and procedures well known in the art.

Compositions for oral use may be in the form of hard gelatin capsules in which the active ingredient is mixed with an inert solid diluent, for example, calcium carbonate, calcium phosphate or kaolin, or as soft gelatin capsules in which the active ingredient is mixed with water or an oil such as peanut oil, liquid paraffin, or olive oil.

Aqueous suspensions generally contain the active ingredient in finely powdered form together with one or more suspending agents, such as sodium carboxymethylcellulose, methylcellulose, hydroxypropylmethylcellulose, sodium alginate, polyvinyl-pyrrolidone, gum tragacanth and gum acacia; dispersing or wetting agents such as lecithin or condensation products of an alkylene oxide with fatty acids (for example polyoxyethylene stearate), or condensation products of ethylene oxide with long chain aliphatic alcohols, for example heptadecaethyleneoxycetanol, or condensation products of ethylene oxide with partial esters derived from fatty acids and a hexitol such as polyoxyethylene sorbitol monooleate, or condensation products of ethylene oxide with long chain aliphatic alcohols, for example

- 39 -

heptadecaethyleneoxycetanol, or condensation products of ethylene oxide with partial esters derived from fatty acids and a hexitol such as polyoxyethylene sorbitol monooleate, or condensation products of ethylene oxide with partial esters derived from fatty acids and hexitol anhydrides, for example polyethylene sorbitan monooleate. The aqueous suspensions
5 may also contain one or more preservatives (such as ethyl or propyl p-hydroxybenzoate, anti-oxidants (such as ascorbic acid), colouring agents, flavouring agents, and/or sweetening agents (such as sucrose, saccharine or aspartame).

Oily suspensions may be formulated by suspending the active ingredient in a vegetable oil (such as arachis oil, olive oil, sesame oil or coconut oil) or in a mineral oil (such as liquid
10 paraffin). The oily suspensions may also contain a thickening agent such as beeswax, hard paraffin or cetyl alcohol. Sweetening agents such as those set out above, and flavouring agents may be added to provide a palatable oral preparation. These compositions may be preserved by the addition of an anti-oxidant such as ascorbic acid.

Dispersible powders and granules suitable for preparation of an aqueous suspension by
15 the addition of water generally contain the active ingredient together with a dispersing or wetting agent, suspending agent and one or more preservatives. Suitable dispersing or wetting agents and suspending agents are exemplified by those already mentioned above. Additional excipients such as sweetening, flavouring and colouring agents, may also be present.

20 The pharmaceutical compositions of the invention may also be in the form of oil-in-water emulsions. The oily phase may be a vegetable oil, such as olive oil or arachis oil, or a mineral oil, such as for example liquid paraffin or a mixture of any of these. Suitable emulsifying agents may be, for example, naturally-occurring gums such as gum acacia or gum tragacanth, naturally-occurring phosphatides such as soya bean, lecithin, an esters or partial
25 esters derived from fatty acids and hexitol anhydrides (for example sorbitan monooleate) and condensation products of the said partial esters with ethylene oxide such as polyoxyethylene sorbitan monooleate. The emulsions may also contain sweetening, flavouring and preservative agents.

Syrups and elixirs may be formulated with sweetening agents such as glycerol, propylene
30 glycol, sorbitol, aspartame or sucrose, and may also contain a demulcent, preservative, flavouring and/or colouring agent.

The pharmaceutical compositions may also be in the form of a sterile injectable aqueous or oily suspension, which may be formulated according to known procedures using one or

- 40 -

more of the appropriate dispersing or wetting agents and suspending agents, which have been mentioned above. A sterile injectable preparation may also be a sterile injectable solution or suspension in a non-toxic parenterally-acceptable diluent or solvent, for example a solution in 1,3-butanediol.

5 Suppository formulations may be prepared by mixing the active ingredient with a suitable non-irritating excipient which is solid at ordinary temperatures but liquid at the rectal temperature and will therefore melt in the rectum to release the drug. Suitable excipients include, for example, cocoa butter and polyethylene glycols.

Topical formulations, such as creams, ointments, gels and aqueous or oily solutions or
10 suspensions, may generally be obtained by formulating an active ingredient with a conventional, topically acceptable, vehicle or diluent using conventional procedure well known in the art.

Compositions for administration by insufflation may be in the form of a finely divided powder containing particles of average diameter of, for example, 30 μ or much less, the
15 powder itself comprising either active ingredient alone or diluted with one or more physiologically acceptable carriers such as lactose. The powder for insufflation is then conveniently retained in a capsule containing, for example, 1 to 50mg of active ingredient for use with a turbo-inhaler device, such as is used for insufflation of the known agent sodium cromoglycate.

20 Compositions for administration by inhalation may be in the form of a conventional pressurised aerosol arranged to dispense the active ingredient either as an aerosol containing finely divided solid or liquid droplets. Conventional aerosol propellants such as volatile fluorinated hydrocarbons or hydrocarbons may be used and the aerosol device is conveniently arranged to dispense a metered quantity of active ingredient.

25 For further information on Formulation the reader is referred to Chapter 25.2 in Volume 5 of Comprehensive Medicinal Chemistry (Corwin Hansch; Chairman of Editorial Board), Pergamon Press 1990.

The amount of active ingredient that is combined with one or more excipients to produce a single dosage form will necessarily vary depending upon the host treated and the particular
30 route of administration. For example, a formulation intended for oral administration to humans will generally contain, for example, from 0.5 mg to 2 g of active agent compounded with an appropriate and convenient amount of excipients which may vary from about 5 to about 98 percent by weight of the total composition. Dosage unit forms will generally contain

- 41 -

about 1 mg to about 500 mg of an active ingredient. For further information on Routes of Administration and Dosage Regimes the reader is referred to Chapter 25.3 in Volume 5 of Comprehensive Medicinal Chemistry (Corwin Hansch; Chairman of Editorial Board), Pergamon Press 1990.

5 The size of the dose for therapeutic or prophylactic purposes of a compound will naturally vary according to the nature and severity of the conditions, the age and sex of the animal or patient and the route of administration, according to well known principles of medicine.

In using a compound for therapeutic or prophylactic purposes it will generally be
10 administered so that a daily dose in the range, for example, 0.5 mg to 75 mg per kg body weight is received, given if required in divided doses. In general lower doses will be administered when a parenteral route is employed. Thus, for example, for intravenous administration, a dose in the range, for example, 0.5 mg to 30 mg per kg body weight will generally be used. Similarly, for administration by inhalation, a dose in the range, for
15 example, 0.5 mg to 25 mg per kg body weight will be used. Oral administration is however preferred.

It will be obvious to those skilled in the art to which the invention pertains, that various changes and modifications may be made without departing from the scope of the invention defined by the claims.

20

EXAMPLES

Example 1

Identification of translocation breakpoints

A male patient having a BMI of 48 was karyotyped and abnormal chromosomes were
25 identified. It was found that the abnormal chromosomes had the translocation products between chromosome 4q22 and chromosome 15q22. The patient's sister and mother also had abnormally high BMI, and were also karyotyped (Fig. 1). It was subsequently found that the sister and mother were also carrying the balanced translocation. The breakpoints of each chromosome were further mapped with fish probes and southern blots and the breakpoints on
30 each chromosome were sequenced. Using bacterial artificial chromosomes (BAC) clones spanning both breakpoints, the breakpoints were sequenced and assembled with the help of the public sequence to generate ordered contigs of the breakpoints region of chromosome 4

- 42 -

and 15. ROR α 1 on chromosome 15 was found to be affected since the breakpoint is within the first intron of isoform 1 resulting in the destruction of one copy of ROR α 1.

Example 2

Gene expression analysis

5 Preadipocytes were isolated from the affected boy and girl and used to study adipocyte function. Gene expression was monitored for all genes mapped to the vicinity of the breakpoints both at the preadipocyte stage and after induction into mature adipocytes. The isolated adipocytes showed an altered phenotype compared with normal adipocytes both in general morphology, function and gene expression. Lipid droplets were incorporated as tiny
10 droplets instead of the large ones normally seen. UNC5C was highly upregulated in the preadipocytes both before and after induction into mature adipocytes when compared with an adipocyte control cell line.

Two differently sized UNC5C transcripts could be detected in preadipocytes isolated from the patients. These were cloned and sequenced and shown to be a fusion between
15 ROR α 1 exon 1 and UNC5C exon 2 giving rise to a novel protein. Both encoded a fusion transcript with the structure ROR α 1-UNC5C fusion (SEQ ID NO: 1 and SEQ ID NO: 3). The shorter of the two (SEQ ID NO: 3) is a splice version using alternative splice sites within exons 11 and 12 of UNC5C.

Patient Material

20 All three patients from the family were carriers of the translocation. The family consisted of affected mother and two affected siblings, the father is normal build. Lymphoblastoid cell lines were established according to standard procedures.

FISH Probes

For yeast artificial chromosome (YAC) screening, YACs from chromosome 4 and 15
25 from the CEPH human YAC library was used (YAC clones chromosome 4: 244C10, 676B9, 767G2, 783H11, 840F2, 854F4, 858A2, 871A1, 881A12 and chromosome 15: 667B2, 798C9, 801B4, 809G4, 827C7, 848E5, 854A11, 895G6, 897E9, 934G1, 937E4, 951F3). STS markers contained in the spanning YAC clones was used for PCR screening of bacterial artificial chromosome (BAC) libraries.

30 The CEPH Human BAC library was screened using STS markers D4S2407 and D4S1559 from YAC clone 858A2 on chromosome 4 and WI-7454 from YAC clone 798C9 on chromosome 15. STS marker AFMa244wc9 from chromosome 15 was used to screen BAC Human II library from Genome Systems, Inc.

- 43 -

Restriction digestion with BamHI and HindIII was used to assemble a contig of the BAC clones and flanking clones from each contig were selected for BAC end walk. BAC end sequencing was performed on whole BACs purified with Qiagen Large Construct Purification kit (Qiagen) using sp6 and T7 primers.

- 5 The end sequences of the BACs were used to screen the *nr* (non-redundant) and *htgs* (high throughput genomic sequences) databases at the NCBI/NIH server using the BLASTN program. Positive BAC hits were ordered from BACPAC Resources (Children's Hospital Oakland) and used as FISH probes.

- To further map the breakpoints, sequences of approximately 6 kb from the spanning
10 BACs RP11-402L11 (AC093832) and RP11-90A19 (AC079068) was derived by long range PCR and used as probes for FISH.

Mini FISH probes used for chromosome 4:

15

Forward primer	Reverse primer
SEQ ID NO: 15	SEQ ID NO: 16
SEQ ID NO: 17	SEQ ID NO: 18
SEQ ID NO: 19	SEQ ID NO: 20
SEQ ID NO: 21	SEQ ID NO: 22
SEQ ID NO: 23	SEQ ID NO: 24
SEQ ID NO: 25	SEQ ID NO: 26

Mini FISH probes used for chromosome 15:

Forward primer	Reverse primer
SEQ ID NO: 27	SEQ ID NO: 28
SEQ ID NO: 29	SEQ ID NO: 30
SEQ ID NO: 31	SEQ ID NO: 32
SEQ ID NO: 33	SEQ ID NO: 34
SEQ ID NO: 35	SEQ ID NO: 36

- 44 -

SEQ ID NO: 37	SEQ ID NO: 38
SEQ ID NO: 39	SEQ ID NO: 40
SEQ ID NO: 41	SEQ ID NO: 42
SEQ ID NO: 43	SEQ ID NO: 44

FISH analysis

FISH was performed on metaphase chromosome spreads from the boy which is carrying the translocation. Metaphase spreads were prepared from lymphocytes according to standard methods.

5

Southern blot

The probes used in southern blot were amplified by PCR from BAC clones RP11-402L11 (chromosome 4) and RP11-90A19 (chromosome 15) and labelled by random priming.

The probes used for chromosome 4 were:

10

Forward primer	Reverse primer
SEQ ID NO: 45	SEQ ID NO: 46
SEQ ID NO: 47	SEQ ID NO: 48
SEQ ID NO: 49	SEQ ID NO: 50
SEQ ID NO: 51	SEQ ID NO: 52
SEQ ID NO: 53	SEQ ID NO: 54

The probes used for chromosome 15 were:

Forward primer	Reverse primer
SEQ ID NO: 55	SEQ ID NO: 56
SEQ ID NO: 57	SEQ ID NO: 58
SEQ ID NO: 59	SEQ ID NO: 60
SEQ ID NO: 61	SEQ ID NO: 62

- 45 -

PCR and Sequencing of the Breakpoints

The breakpoints could be bridged by long range PCR with chromosome derivate specific primers. For the chromosome 4 derivate the chromosome 4 primer SEQ ID NO: 63 and the chromosome 15 primer SEQ ID NO: 64 was used, and for the chromosome 15 derivate the
5 chromosome 15 primer SEQ ID NO: 65 and the chromosome 4 primer SEQ ID NO: 66 was used.

The breakpoint on the chromosome 4 derivate PCR product could be sequenced using primer SEQ ID NO: 51 and the chromosome 15 derivate PCR product could be sequenced using primer SEQ ID NO: 52.

10 *RNA isolation*

Total RNA was isolated from cells using Trizole reagent (Gibco) according to the manufacturer's description. The isolated RNA was dissolved in DEPC treated water and concentration was determined by spectrophotometric measurement at 260nm. All samples were also electrophoresed in 1% agarose gel to assess integrity of ribosomal bands.

15 *DNase treatment and cDNA synthesis*

15µg of total RNA was treated with DNase to remove contaminating genomic DNA before cDNA synthesis. DNA-free™ from Ambion was used for this purpose. Shortly; 0.1 volume of 10 times DNase buffer, 6 units of DNase and water up to 30µl was added to each sample. Samples were incubated at 37° for 40 minutes. DNase inactivation reagent was added
20 to each sample and after a 2 minutes incubation at room temperature it was removed by centrifugation.

First strand cDNA synthesis was performed with SuperScript™ First-Strand Synthesis System for RT-PCR (Invitrogen, Life Technologies) according to enclosed protocol. 2µg of DNase treated total RNA was used in each synthesis. Priming of synthesis was accomplished
25 with Oligo(dT)₁₂₋₁₈. For each synthesis reaction a negative control with no enzyme was set up.

PCR and cloning

Full length cDNA of the fusion transcript was obtained with TaqPlus® Precision PCR System (Stratagene). In a total of 50µl, 1 time TaqPlus Precision buffer, 200µM of each dNTP (Applied Biosystems), 300nM forward primer SEQ ID NO: 67, 300nM reverse primer
30 SEQ ID NO: 68 and 2.5U TaqPlus Precision polymerase mixture. cDNA from above was used as template, 0.4µl/reaction (equivalent with 40ng total RNA). The reaction mixture was

- 46 -

then subjected to thermal cycling in an PTC-200 from MJ research. Thermal cycling included one step initial denaturation one minute at 95°C, 35 cycles with 95°C one minute, 58°C one minute, 72°C three minutes and a final extension of ten minutes at 72°. The reaction mixture was then loaded on a 1% Low melt Agarose gel (BioRad). Inspection of the gel
5 showed 2 bands of approximately 2.8 and 3 kb. Both bands were cut out and ligated into pcR2.1-TOPO® (Invitrogen) and sequentially transfected into TOP 10 (Invitrogen).

Both constructs were sequenced using Big Dye chemistry and 3100 Genetic Analyzer(Applied Biosystems).

Quantification

10 cDNA from above was also used for relative quantification on Applied Biosystems 7700.

Quantitative real-time PCR was carried out using SYBR® Green PCR Master Mix in 25µL reactions run in triplicates. Template for the PCR was first strand cDNA from above (amount equivalent of 10ng total RNA). Forward and reverse primers specific for UNC5c was used at a concentration of 400nM. As negative control for amplification of remaining
15 genomic DNA, no enzyme controls from cDNA synthesis was used.

Shotgun library construction

Human Bac clones were cultured over night in LB-media. DNA was prepared using Maxi-plasmid prep from Qiagen (VWR).

5µg purified DNA was fragmented with a nebulizer.

20 DNA was fragmented to 1 to 3 kb fragments, checked on 1% agarose gel.

All according to the manufacturer's description in TOPO ® Shotgun Sub cloning Kit (Invitrogen). DNA was Blunt-end repaired, dephosphorylated and transformed into One Shot competent E.coli, also from the kit.

The transformation reaction was plated on LB plates and incubated over night.

25 A colony picker Q-PIX (Genetix) picked the colonies into 96-well micro plates with LB-media.

One more incubation 37°C over night.

Hydra, a dispenser moved the bacterial suspension to 96-deep well plates and a 34-hour 37°C shake incubation was performed. After centrifugation, continued with DNA preparation.

- 47 -

DNA preparation

Add 50 μ l water, vortex 15 minutes. When complete resuspension add 70 μ l Stet-Tween20 with 5,2mg/ml Rnase and 8,4mg/ml Lysozyme. Denaturate in microwave 1 minute add 335 μ l water, vortex and let stand in ice for 10 minutes. Spin boxes for 30 minutes x 4000rpm. The clear supernatant moves to new 96-well micro plates to be sequenced.

Sequencing

Big Dye chemistry primer T3, primer T7 and ABI PRISM 3700 DNA Analyzer (Applied Biosystems) was used for sequencing clones.

10 *2 x 96 plate*

Big Dye mix	220 μ l (v.2.0)
Water	220 μ l
Buffer x 2,5	660 μ l (200mMTris-HCl ph 9.0, 5mM MgCl ₂)
Primer T7 or T3	7 μ l (100pmol/ μ l)

15 Robot Biomek 2000, loaded the mix and the DNA samples into PCR-plates.

PTC-200, (MJ Research) the thermal cycling instrument;

95°C x 2', (95°C x 10", 50°C x 10", 60°C x 4') x 35 cycles and 8°C.

Followed by centrifugation and ethanol cleaning up step.

15 μ l water was added before running on 3700.

20 *5'RACE of ROR α*

Template used was Marathon Ready cDNA from Clontech (Human Adipocyte #7447(pooled from 11 male/female Caucasians, ages 19-57)).

Nested PCR was used to obtain discrete bands from the PCR reactions. In both reactions following concentrations was used; forward and reverse primer 200nM, dNTP 200 μ M, MgCl

25 2mM, 0.125U/ μ L AmpliTaq Gold (Applied Biosystems) all in one time PCR buffer II (Applied Biosystems).

In the first round, forward primer was AP1 (linker primer, Clontech) (SEQ ID NO: 69).

Reverse primer, situated in the RORalpha sequence (SEQ ID NO: 70). Template 2 μ L

(~0.1ng/ μ L, Marathon Ready cDNA). This reaction mixture was then subjected to thermal

30 cycling in a PTC-200 from MJ research. Thermal cycling included one step initial

- 48 -

denaturation of eight minutes at 95°C, 30 cycles with 94°C fortyfive seconds, 66°C fortyfive seconds, 72°C one minute and a final extension of ten minutes at 72°.

In the second round of PCR, forward primer was AP2 (linker primer, Clontech) (SEQ ID NO: 71). Reverse primer was the same as was used in the first round of PCR (SEQ ID NO: 70).

- 5 Template was 1µL of a twenty times dilution of the first PCR-reaction. This mixture was subjected to another 30 cycles, as above, but with 64 instead of 66°C as annealing temperature.

PCR reactions were loaded on a low melt agarose gel (1% from BioRad). Potentially interesting bands were chosen and cut out from the gel, slices melted at 65° and ligated into 10 pcR2.1-TOPO® (Invitrogen). Ligation mixture was then used to transfect TOP 10 (Invitrogen).

Two clones from each band was then sequenced using Big Dye chemistry and 3100 Genetic Analyzer(Applied Biosystems).

Taqman analysis of RORα5

- 15 Primers where designed with Applied Biosystems program Primer Express 2.0 Forward primer Tm 58,5°C was SEQ ID NO: 72. Reverse primer Tm 58,5°C was SEQ ID NO: 73. Template poly A+RNA from the following human tissues was obtained from Clontech; Bone marrow, brain, kidney, liver, lung, skeletal muscle, small intestine, spleen, testis and fat. CDNA synthesis was made with SuperScript™ First-Strand Synthesis System 20 for RT-PCR, Invitrogen™ according to the protocol supplied by the manufacturer. Real time PCR was performed according to protocol from the manufacturer and run on ABI Prism 7700® Sequence Detection System, Applied Biosystems. Results analysed with Sequence detector V1.7

PCR detection of UNC5C splice variant transcript

- 25 Forward primer was SEQ ID NO: 74 spanning bases 1586-1604 of human UNC5C sequence AF055634. Reverse primer was SEQ ID NO 75 spanning 2386-2367 of transcript AF055634. Template was cDNA from human brain (mRNA from Clontech converted into cDNA using "Superscript™ First-Strand Synthesis System for RT-PCR" (Invitrogen)) according to protocol from the manufacturer. This PCR resulted in 3 bands. The expected one 30 at ~800bp, a very faint one that was slightly shorter and one at ~550bp. The 800 and 550 bands were cut out of the gel and purified using Qiaquick gel extraction kit (Qiagen) and

- 49 -

then sequenced using Big Dye chemistry and 3100 Genetic Analyzer (Applied Biosystems). The 800 band was the expected UNC5c sequence i.e. partial UNC5c spanning bases 1586-2386. The 550bp fragment was UNC5C with a deletion of bases 1873-2078. This deletion is identical to the deletion in the sequence of UNC5Cb with accession number AAC90914.

5 *Bioinformatics*

Sequence assembly and alignment of gene sequences was made using Sequencer software from Gene codes corporation version 4.1.2 on a Macintosh G4.

Primary cell culture

The following products were purchased from BioWhittaker (MD, USA): DMEM F-12, 10 FBM (fibroblast basal medium), human fibroblast growth factor- β (hFGF- β ; 1 μ g/ml), Trypsin EDTA (200 mg/ml and 500 mg/ml), Fungizone (250 μ g/ml). Foetal calf serum (FCS) was provided from Harlan Sera lab (Leicestershire, UK). Pioglitazone (X-46) was from AstraZeneca R&D Molndal. 0.008M Phosphate buffered saline (PBS); insulin and transferrin were obtained from Life 15 Technologies. BSA fraction V, Collagenase (C-6885), IBMX (3-isobutyl-1-methylxanthine), dexamethasone, 0.05M Tris Buffered Saline pH 8.0 were purchased from SIGMA. β -mercaptoethanol and glucose were from Kebo lab. Primaria™ tissue culture flasks were from Falcon. Nile Red was from Molecular Probes.

Biopsies from humans were taken under aseptic conditions and transferred to 20 DMEM/F12 supplemented with 25 μ g/ml Gentamycin and Fungizone (2.5 μ g/ml). The adipose tissue was digested in DMEM F-12 containing albumin (1 mg/ml), collagenase (1 mg/ml), glucose (2 mg/ml), 25 μ g/ml Gentamycin and Fungizone (2.5 μ g/ml) in a shaking waterbath at 37°C for 30-45 min. The cell suspension was then filtered through a nylon mesh (100 μ m) and centrifuged at 650g for 10 minutes. The sedimented cells, preadipocytes, were 25 resuspended and washed twice in DMEM F-12 supplemented with 25 μ g/ml Gentamycin and Fungizone (2.5 μ g/ml). After the last wash the cells were resuspended in FBM (fibroblast medium) supplemented with 10% FCS, 25 μ g/ml Gentamycin and Fungizone (2.5 μ g/ml, 0.1% 0.1M mercaptoethanol, 10 mg/ml insulin, 10 mg/ml transferrin and 0.001 μ g/ml hFGF- β . The cells were seeded and incubated in small Primaria culture flasks (25 cm²) over night 30 followed by washing with PBS and culture continued until the cells reached confluence. Thereafter the cells were washed in PBS (without Mg⁺⁺ and Ca⁺⁺), treated with trypsin EDTA for a few minutes, resuspended and split in fresh FBM and seeded into new small culture

- 50 -

flasks. The medium was changed every day. The new cell lines were named PSc020108 (male) and Asc020108 (female).

The fraction with the mature adipocytes (after protease treatment) was frozen in -80°C for further RNA preparation.

5 *Differentiation*

The undifferentiated preadipocytes were seeded into small cell culture flasks and cultured in DMEM/F12 supplemented with 10% FCS, 1% PEST, 10 µg/ml insulin, 10 µg/ml transferrin and 0.1 mM β-mercaptoethanol (referred to as standard medium). When the cells have reached confluence the differentiation was induced by adding standard medium

- 10 supplemented with 0.25µM dexamethasone and 0.5 mM IBMX (referred to as induction medium 1) for three days. To maintain and extend the differentiation process the medium was changed to standard medium containing 10 µM X-46 and 0.25µM Dexamethasone (referred to as induction medium 2) for 2 days. This process was repeated once and after that the culture process was continued only with standard medium supplemented with 0.25µM
- 15 dexamethasone (referred to as maintenance medium1).

The lipid accumulation in the adipocytes was confirmed by Nile Red staining.

Staining cells with Nile Red

- Nile red is a benzophenoxazone dye. In older chemical and histochemical literature the dye sometimes referred to as nileblue A-oxazone. Nile red is poorly soluble in water but does
- 20 dissolve in a wide variety of organic solvents.

A 1mg/ml stock solution of Nile red in acetone were prepared and stored protected from light. Staining could be carried out on either fixed (1.5% glutaraldehyde in PBS, for 5 min) or unfixed cells. Attached unfixed cultured cells were covered with PBS. The dye was then added directly to the flask to effect a 1:100 dilution to yield a final concentration of 10 µg/ml.

- 25 The preparation was incubated for a minimum of 5 min. For fluorescence microscopy excess dye was removed by briefly rinsing the cells with PBS. The suspension medium must not contain serum or albumin since they could act as a sink to draw Nile red out of the cells. Dye in the medium does not interfere with the observations since Nile red fluorescence is quenched in water. Nile red is not toxic for the cells so it is possible to continue the culture
- 30 after staining when unfixed cells are used.

Morphology and Fluorescence microscopy

The differentiation of adipocyte precursors into mature adipocytes was assessed by conventional microscopy at a 100-200-fold magnification. Cells were considered

- 51 -

differentiated when their cytoplasm was filled with lipid droplets, clearly seen in phase contrast. The fluorescence microscopy studies was carried out using a Zeiss axiovert 25 inverted microscope, equipped with a HBO 50 illuminator and with a CP-Achromat 10x/0.25 ph1/Var1 objective lens and the excitation wavelength was 485 nm (filter set 487909).

5 The stained cells were photographed in color using the AxioVision system from Carl Zeiss. The pictures were processed in Adobe PhotoShop 6.0.

- 52 -

CLAIMS

1. An isolated nucleic acid molecule comprising a nucleotide sequence having at least 65% identity to a degenerate variant of SEQ ID NO: 1, SEQ ID NO: 3 or SEQ ID NO: 7.

2. The isolated nucleic acid molecule of claim 1, wherein the nucleic acid molecule comprises the nucleotide sequence of SEQ ID NO: 1, SEQ ID NO: 3 or SEQ ID NO: 7.

3. An isolated nucleic acid molecule comprising the complement of SEQ ID NO: 1, SEQ ID NO: 3 or SEQ ID NO: 7.

4. An isolated nucleic acid molecule encoding a ROR α 1-UNC5C polypeptide, said nucleic acid molecule comprising a nucleotide sequence having at least 65% identity to a degenerate variant of SEQ ID NO: 1 or SEQ ID NO: 3.

5. An isolated nucleic acid molecule encoding a ROR α 5 polypeptide, said nucleic acid molecule comprising a nucleotide sequence having at least 65% identity to a degenerate variant of SEQ ID NO: 7.

6. A vector comprising the nucleic acid molecule of any of claims 1-5.

7. A host cell comprising the vector of claim 6.

8. A purified fusion polypeptide of the ROR α 1-UNC5C polypeptide comprising an amino acid sequence having at least 65% identity to the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4.

9. The purified polypeptide of claim 8, wherein the amino acid sequence comprises the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4.

10. A purified polypeptide of the ROR α 5 polypeptide comprising an amino acid sequence having at least 65% identity to the amino acid sequence of SEQ ID NO: 8.

- 53 -

11. The purified polypeptide of claim 10, wherein the amino acid sequence comprises the amino acid sequence of SEQ ID NO: 8.

12. A method for producing a protein comprising:

- 5 a) culturing the host cell of claim 7 under conditions whereby the protein is produced, and
- b) recovering the protein from the host cell culture.

13. A method for detecting a polynucleotide which encodes a protein comprising the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4 or SEQ ID NO: 8 in a biological sample comprising the steps of:

- a) hybridizing a complement of the polynucleotide sequence which encodes SEQ ID NO: 2, SEQ ID NO:4 or SEQ ID NO: 8 to a nucleic acid material of a biological sample, thereby forming a hybridization complex; and
- 15 b) detecting the hybridization complex, wherein the presence of the complex correlates with the presence of a polynucleotide encoding a protein in the biological sample.

14. A method for detecting the presence of an obesity susceptibility gene comprising:

- determining the identity of a gene located 20 cM, or less, adjacent to the
- 20 translocation breakpoint between chromosome 4 at cytoband 4q22.3 and chromosome 15 at cytoband 15q22.2; and

determining if the gene has aberrant gene expression as a consequence of the translocation event.

25 15. A method for detecting the presence of a translocation junction between chromosome 4 at cytoband 4q22.3 and chromosome 15 at cytoband 15q22.2 comprising analysing a sample of DNA from an individual for the presence of the translocation junction.

16. A method for identifying a test compound that modulates the expression of an obesity susceptibility gene identified in claim 14 comprising:

- 30 contacting a cell capable of expressing the susceptibility gene with a test compound; and
- determining the level of expression of the obesity susceptibility gene in the presence of the test compound, wherein a decrease or an increase in expression of a obesity susceptibility

- 54 -

gene, as compared to the level of expression of an obesity susceptibility gene in the absence of the compound, is indicative that the test compound modulates the expression of the obesity susceptibility gene.

5 17. A method of identifying a test compound that modulates the activity of an obesity protein encoded by the obesity susceptibility gene identified in claim 14, comprising:

contacting the obesity protein with a test compound; and

determining the level of activity of the obesity protein in the presence of the compound,

wherein a decrease or an increase in obesity protein activity, as compared to the level of
10 activity of the obesity protein in the absence of the compound, is indicative that the test compound modulates obesity protein activity.

18. A method of treating a subject having obesity comprising administering an effective amount of the compound identified in claims 16 or 17.

15

19. Use of a compound identified in claims 16 or 17 in the preparation of a medicament for the treatment of obesity.

20. A pharmaceutical composition comprising a compound identified in claims 16 or 17, and
20 a pharmaceutically acceptable adjuvant, diluent or carrier.

21. A method of making a pharmaceutical composition comprising:

contacting a cell capable of expressing an obesity susceptibility gene with a test compound;

25 determining the level of expression of the obesity susceptibility gene in the presence of the test compound, wherein a decrease in expression of a obesity susceptibility gene, as compared to the level of expression of a obesity susceptibility gene in the absence of the compound, is indicative that the test compound decreases obesity susceptibility gene expression; and

30 formulating the test compound that decreases obesity susceptibility gene expression into a pharmaceutical composition.

- 55 -

22. A method of making a pharmaceutical composition comprising:
contacting an obesity protein with a test compound;
determining the level of activity of the obesity susceptibility protein in the presence of the compound, wherein a decrease in obesity protein activity, as compared to the level of activity
5 of the obesity protein in the absence of the compound, is indicative that the test compound decreases obesity protein activity; and
formulating the test compound that decreases obesity protein activity into a pharmaceutical composition.
- 10 23. A method for determining if an obesity susceptibility gene identified in claim 14 has an altered level of gene expression comprising:
comparing the level of obesity gene expression in a cell from a patient having obesity with a control cell, and
determining the level of expression of the obesity susceptibility gene in both cells,
15 wherein a decrease or an increase in expression of the obesity susceptibility gene, as compared to the level of expression of the obesity susceptibility gene in the control cell, indicates that the obesity susceptibility gene has altered gene expression.
24. A method of diagnosing obesity, or a susceptibility thereto in a subject, the method
20 comprising:
determining the level of mRNA of UNC5C; and
comparing the level of mRNA of UNC5C in the sample with a control, wherein an increase in the level of UNC5C in the sample compared to the control indicates that the subject has obesity, or a susceptibility thereto.
- 25 25. A method of diagnosing obesity or a susceptibility thereto in a subject, the method comprising:
determining the level of aUNC5C protein in a sample from a subject; and
comparing the level of UNC5C in the sample with a control, wherein an increase in the
30 level of the protein in the sample compared to the control indicates that the subject has obesity, or a susceptibility thereto.

- 56 -

26. A method of diagnosing obesity, or a susceptibility thereto in a subject, the method comprising analysing for the presence of ROR1 α 1-UNC5C mRNA, wherein the presence of the mRNA indicates that the subject has obesity, or a susceptibility thereto.
- 5 27. A method of diagnosing obesity or a susceptibility thereto in a subject, the method comprising analysing for the presence of the ROR1 α 1-UNC5C fusion polypeptide, wherein the presence of the polypeptide indicates that the subject has obesity, or a susceptibility thereto.

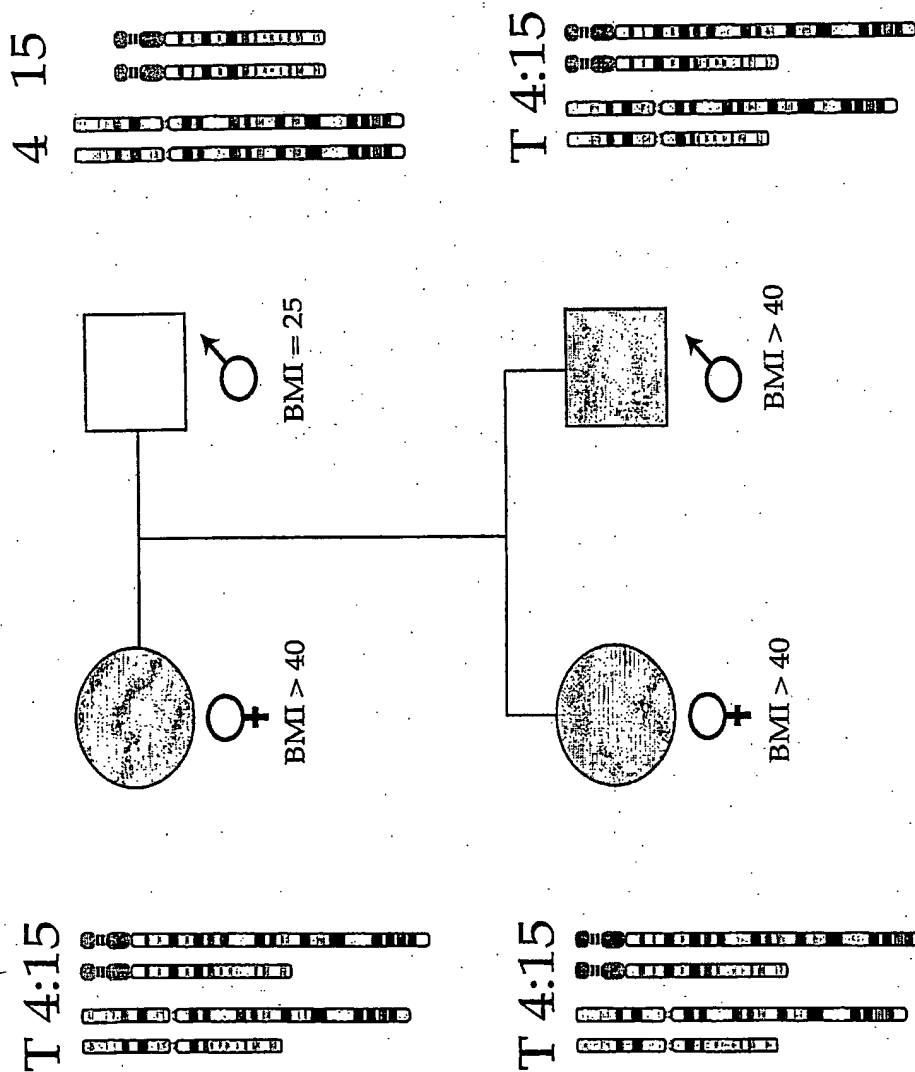


FIGURE 1

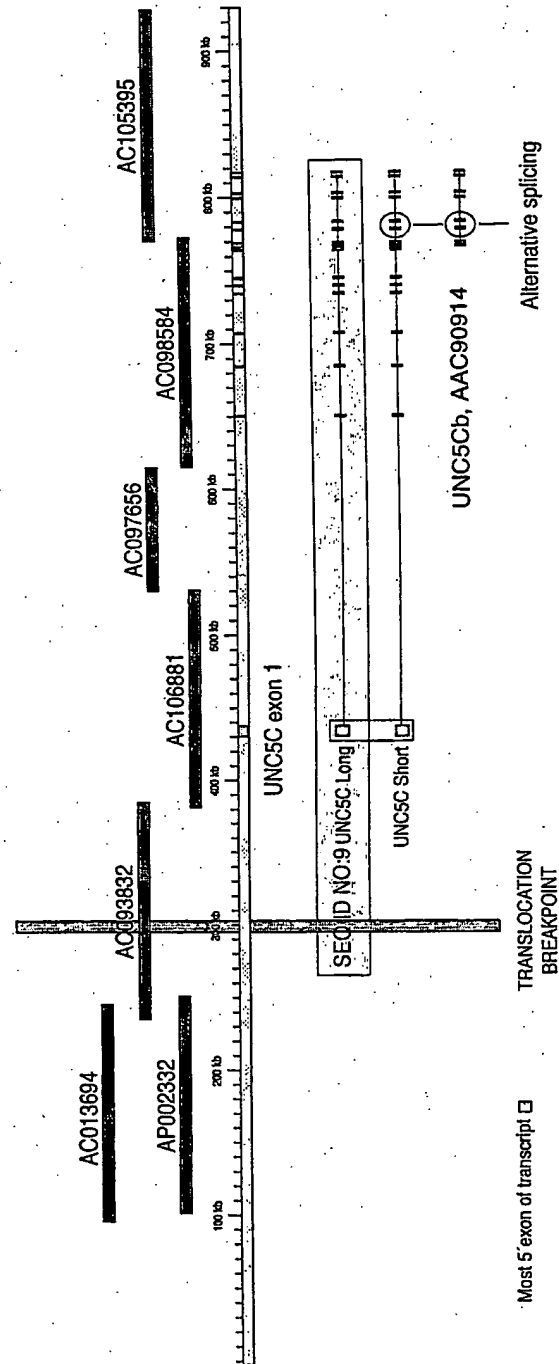


FIGURE 2

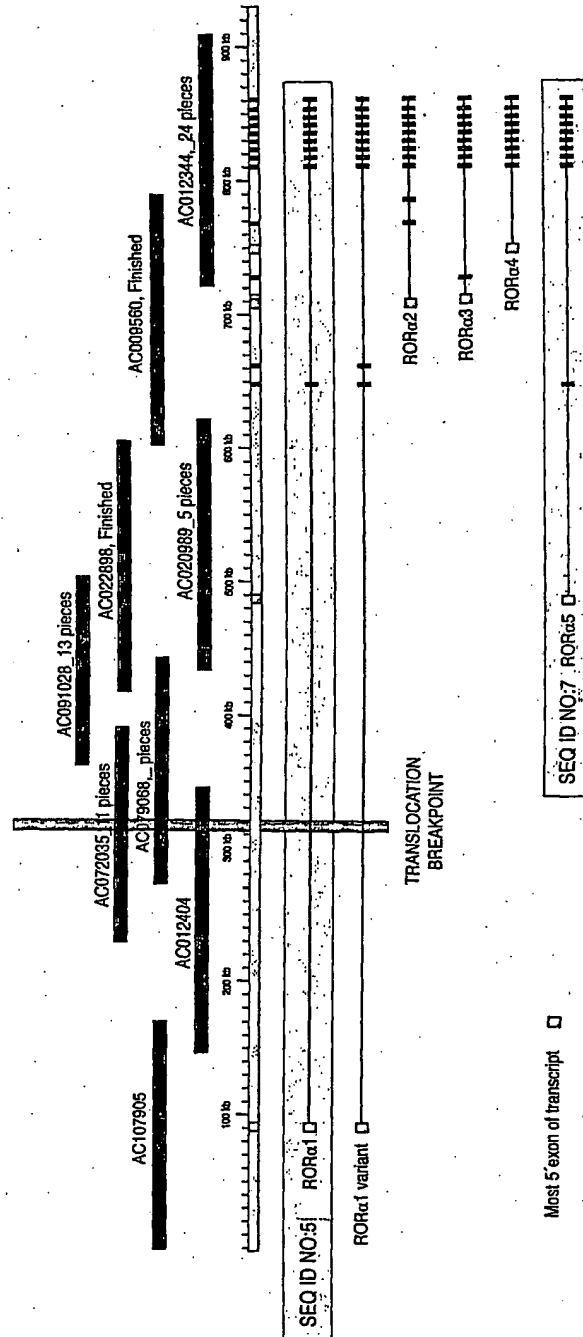


FIGURE 3

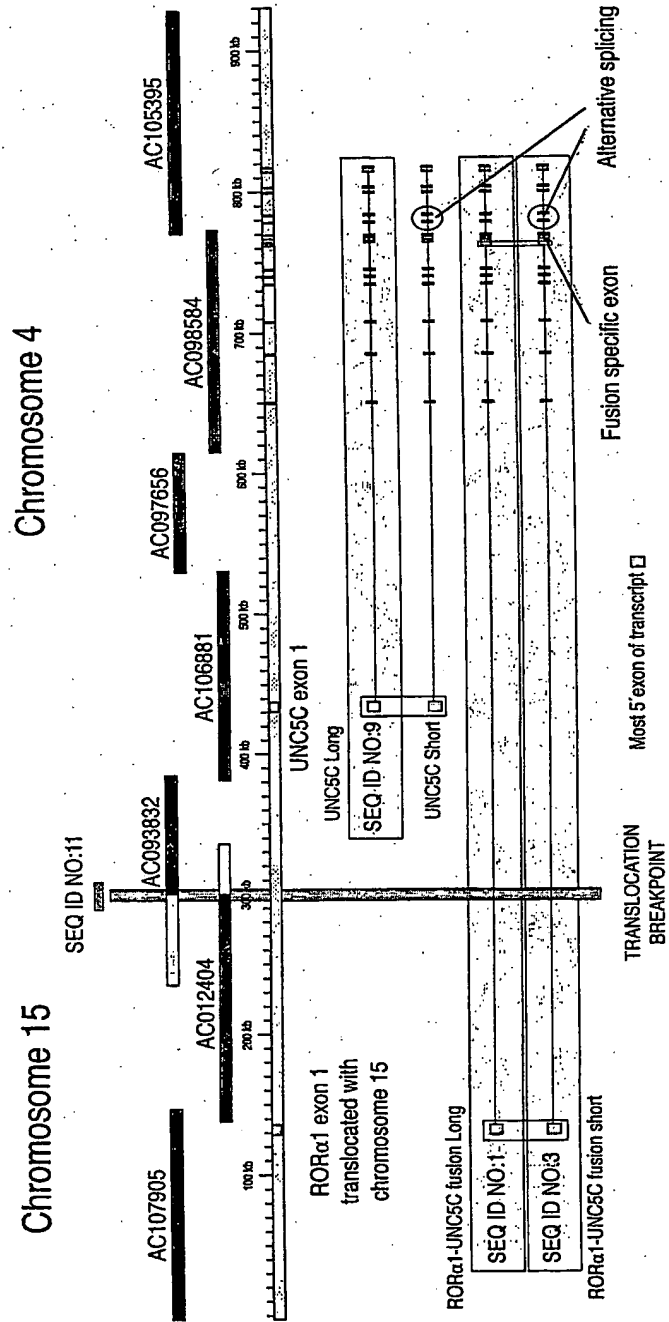


FIGURE 4

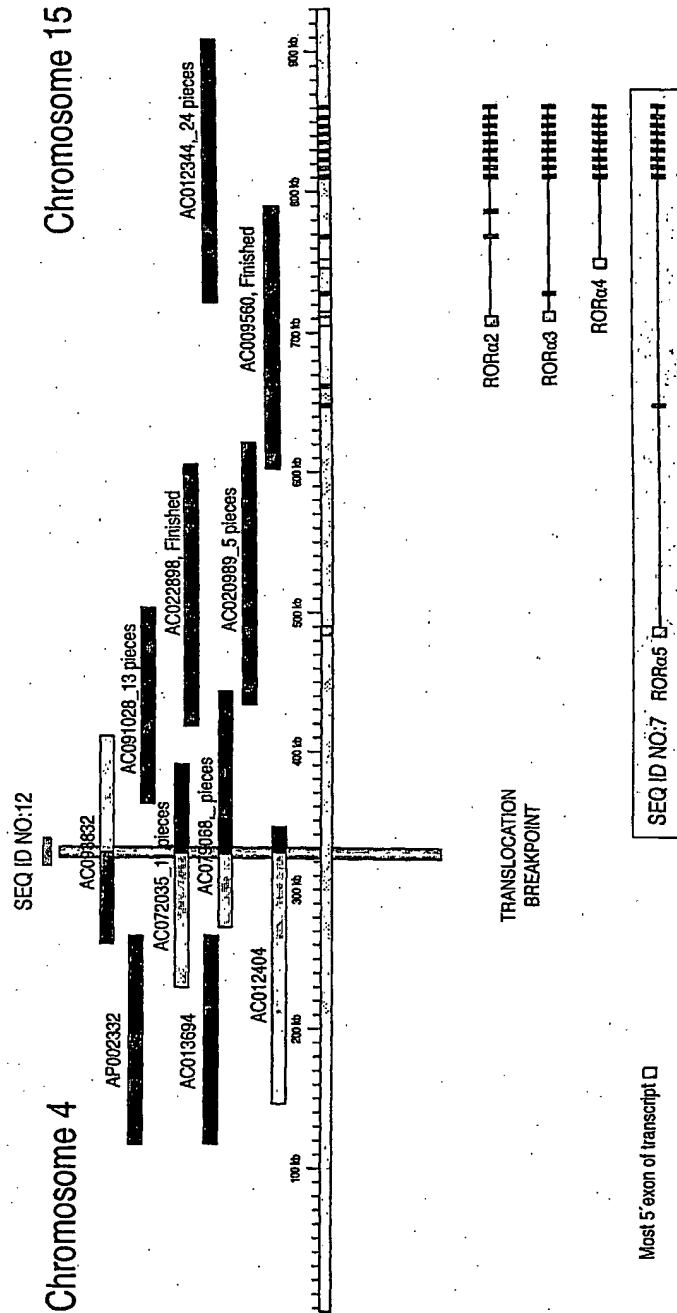


FIGURE 5

6/8

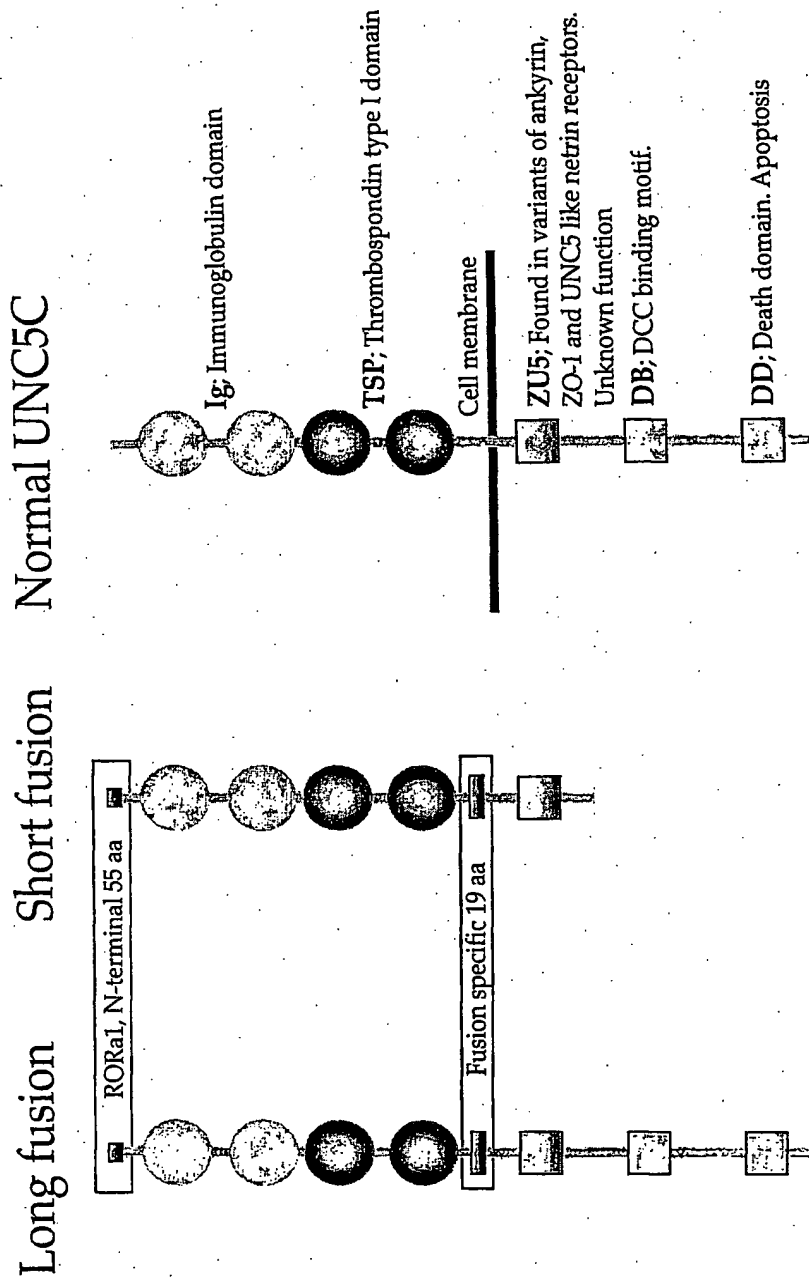


FIGURE 6

7/8

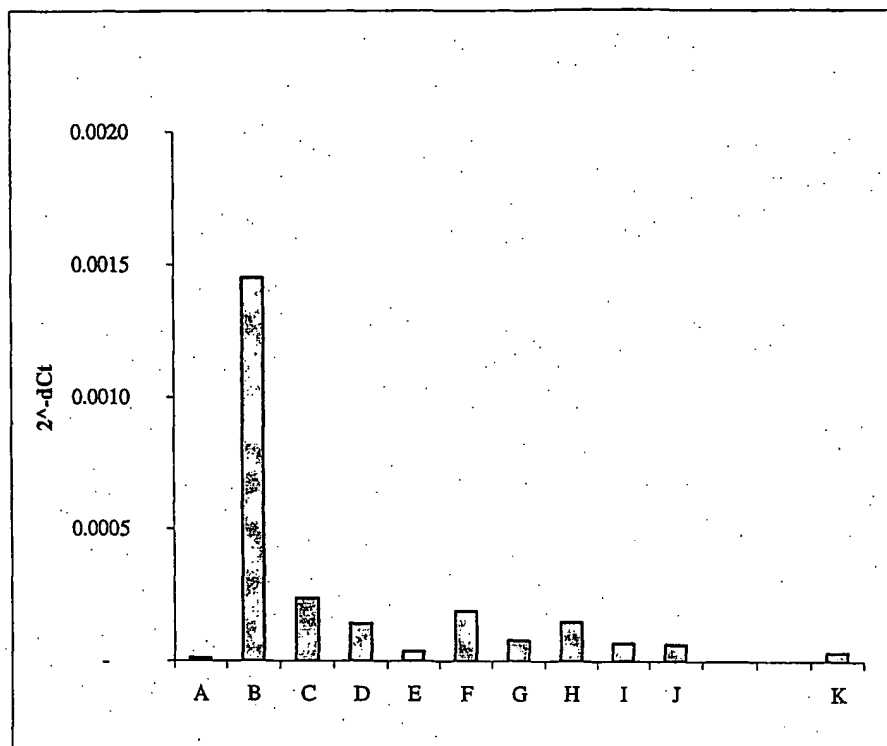
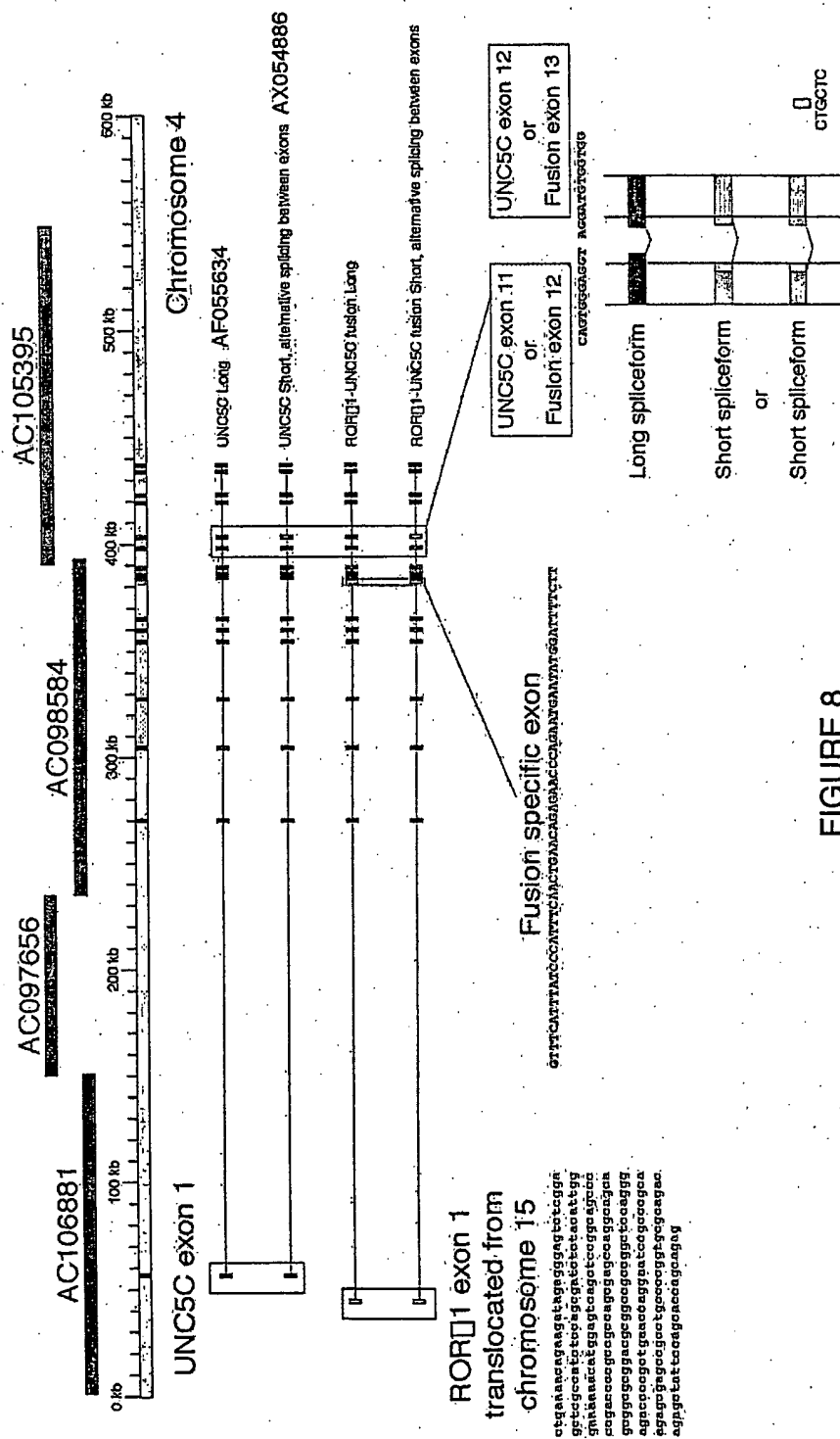


FIGURE 7



WO 2004/083371

PCT/GB2004/001524

10/549460

- 1 -

SEQUENCE LISTING

<110> AstraZeneca AB

5 <120> Molecules

<130> CCH/100852

<150> GB 0306185

10 <151> 2003-03-19

<160> 1

<170> PatentIn version 3.2

15

<210> 1

<211> 2986

<212> DNA

<213> Homo sapiens

20

<400> 1

ctgaaaaacag	aagatagaga	ggagtctcgg	agctcgccat	ctccagcgat	ctctacattg	60
ggaaaaaaca	tggagtcagc	tccggcagcc	cccgaacccg	ccgccagcga	gccaggcagc	120
agcggcgcg	agcggcgccg	cgggtccagg	gagaccccg	tgaaccagga	atccgcccgc	180
aagagcgagc	cgcctgcccc	ggtgcgcaga	cagagctatt	ccagcaccag	cagagatgat	240
gacttttttc	atgaactccc	agaaaactttt	ccttctgac	cacctgagcc	tctgccacat	300
ttccttattg	agcctgaaga	agcttatatt	gtgaagaata	agcctgtgaa	cctgtactgt	360
aaagcaagcc	ctgccaccca	gatctatttc	aagtgtata	gtgaatgggt	tcatcagaag	420
gaccacatag	tagatgaaag	agtagatgaa	acttcgggtc	tcattgtccg	ggaagtgagc	480
attgagattt	cgcgcagca	agtgggaaga	ctctttggac	ctgaagatta	ctggtgccag	540
tgtgtggcct	ggagctccgc	gggtaccaca	aagagccgga	aggcgtatgt	gogcattgca	600
tatctacgga	agacatttga	gcaggaaacc	ctaggaaagg	aagtgtcttt	ggaacaggaa	660
gtcttactcc	agtgtcgacc	acctgaagg	atccagtggt	ctgagggtgga	atggttgaaa	720
aatgaagaca	taattgatcc	cgttgaagat	cggaattttt	atattactat	tgatcacaac	780
ctcatcataa	agcaggcccg	actctctgat	actgcaaatt	acacctgtgt	tgccaaaaac	840
attgttgcca	agaggaaaag	tacaactgcc	actgtcatag	tctatgtcaa	cgggtggctgg	900
tccacctgga	cggagtggtc	tgtgtgtaac	agccgctgtg	gacgagggtg	tcagaaacgt	960
acaaggactt	gtaccaaccc	ggcaccactc	aatgggggtg	ccttctgtga	agggcagagt	1020
gtgcagaaaa	tagcctgtac	tacgttatgc	ccagtggatg	gcagggtggac	gccatggagc	1080
aagtgtgcta	cttgtggaac	tgagtgcacc	cactggcgca	ggagggtgag	cacggcgcca	1140
gcccccaaga	atggaggcaa	ggactgcgac	ggcctcgtct	tgcaatccaa	gaactgcact	1200
gatgggcttt	gcatgcagag	tttcatttat	cccatttcaa	ctgaacagag	aaccagaat	1260
gaatatggat	ttctctctgc	tcctgattca	gatgatgttg	ctctctatgt	tgggattgtg	1320
atagcagtga	tcgtttgcct	ggcgatctct	gtagttgttg	ccttgtttgt	gtatcggaag	1380
aatcatcgty	acttttagtc	agatattatt	gactcttcgg	cactcaatgg	gggctttcag	1440
cctgtgaaca	tcaaggcagc	aagacaagat	ctgctggctg	tacccccaga	cctcacgtca	1500
gctgcagcca	tgtagagagg	acctgtctat	gccctgcatt	acgtctcaga	caaaatccca	1560
atgaccaact	ctccaattct	ggatccactg	cccaacctga	aaatcaaagt	gtacaacacc	1620
tcagggtgctg	tcacccccca	agatgacctc	tctgagttta	cgtccaagct	gtcccctcag	1680
atgaccagtg	cgttgttgga	gaatgaagcc	ctcagcctga	agaaccagag	tctagcaagg	1740
cagactgata	catcctgtac	cgcatttggc	agcttcaact	cactgggagg	tcaccttatt	1800
gttcccaatt	caggagtcag	cttgctgatt	cccgctgggg	ccattcccca	agggagagtc	1860
tacgaaatgt	atgtgactgt	acacaggaaa	gaaactatga	ggccacccat	ggatgactct	1920
cagacacttt	tgacccctgt	ggtgagctgt	gggccccccg	gagctctgct	caccgccttc	1980
gtcgtcctca	ctatgcatca	ctgcgcagac	cccaataacc	aggactggaa	aatactgctc	2040
aagaaccagg	cagcacaggg	acagtgggag	gatgtgggtg	tggtcgggga	ggaaaacttc	2100
accacccctt	gctacattca	gctggatgca	gaggcctgcc	acatcttcac	agagaacctc	2160
agcacctacg	ccctggtagg	acatttcacc	accaaagcgg	ctgcaaagcg	cctcaagctg	2220
gccatctttg	ggccctctgt	ctgctcctcg	ctggagtaca	gcacccaggt	ctactgtctg	2280
gatgacaccc	aggatgcctt	gaaggaaatt	ttacatcttg	agagacagac	gggaggacag	2340
ctcctagaag	aacctaaagg	tcttcatttt	aaaggcagca	cccacaacct	gcgctgtgca	2400
attcacgata	tcgcccattc	cctctggaag	agcaaatgtc	tggctaaata	tcaggaaatt	2460
ccattttacc	atgtttggag	tggaatctca	agaaacctgc	actgcacctt	cactctggaa	2520
agatttagcc	tgaacacagt	ggagctgggt	tgcaaatctt	gtgtcgggca	gggtggaagga	2580
gaagggcaga	ctctccagct	caactgcacc	gtgtcagagg	aacctactgg	catcgatttg	2640
ccgctgctgg	atcctgcgaa	caccatcacc	acgggtcacg	ggccagtgct	tttcagcatc	2700
cctctcccta	tccggcagaa	gctctgtagc	agcctggatg	ccccccagac	gagaggccat	2760
gactggagga	tgctggccca	taagctgaac	ctggacaggt	acttgaatta	ctttgccacc	2820

- 2 -

aaatccagcc caactggcgt aatcctggat ctttggaag cacagaactt cccagatgga 2880
 aacctgagca tgctggcagc tgccttgga gaaatgggaa gacatgaaac ggtggtgtcc 2940
 ttagcagcag aagggcagta ttaaccacca tgctggaagg ggaaat 2986

5

<210> 2
 <211> 964
 <212> PRT
 <213> Homo sapiens

10

<400> 2
 Met Glu Ser Ala Pro Ala Ala Pro Asp Pro Ala Ala Ser Glu Pro Gly
 1 5 10 15

15

Ser Ser Gly Ala Asp Ala Ala Ala Gly Ser Arg Glu Thr Pro Leu Asn
 20 25 30

Gln Glu Ser Ala Arg Lys Ser Glu Pro Pro Ala Pro Val Arg Arg Gln
 35 40 45

20

Ser Tyr Ser Ser Thr Ser Arg Asp Asp Asp Phe Phe His Glu Leu Pro
 50 55 60

25

Glu Thr Phe Pro Ser Asp Pro Pro Glu Pro Leu Pro His Phe Leu Ile
 65 70 75 80

Glu Pro Glu Glu Ala Tyr Ile Val Lys Asn Lys Pro Val Asn Leu Tyr
 85 90 95

30

Cys Lys Ala Ser Pro Ala Thr Gln Ile Tyr Phe Lys Cys Asn Ser Glu
 100 105 110

Trp Val His Gln Lys Asp His Ile Val Asp Glu Arg Val Asp Glu Thr
 115 120 125

35

Ser Gly Leu Ile Val Arg Glu Val Ser Ile Glu Ile Ser Arg Gln Gln
 130 135 140

40

Val Glu Glu Leu Phe Gly Pro Glu Asp Tyr Trp Cys Gln Cys Val Ala
 145 150 155 160

Trp Ser Ser Ala Gly Thr Thr Lys Ser Arg Lys Ala Tyr Val Arg Ile
 165 170 175

45

Ala Tyr Leu Arg Lys Thr Phe Glu Gln Glu Pro Leu Gly Lys Glu Val
 180 185 190

Ser Leu Glu Gln Glu Val Leu Leu Gln Cys Arg Pro Pro Glu Gly Ile
 195 200 205

50

Pro Val Ala Glu Val Glu Trp Leu Lys Asn Glu Asp Ile Ile Asp Pro
 210 215 220

55

Val Glu Asp Arg Asn Phe Tyr Ile Thr Ile Asp His Asn Leu Ile Ile
 225 230 235 240

Lys Gln Ala Arg Leu Ser Asp Thr Ala Asn Tyr Thr Cys Val Ala Lys
 245 250 255

60

Asn Ile Val Ala Lys Arg Lys Ser Thr Thr Ala Thr Val Ile Val Tyr
 260 265 270

Val Asn Gly Gly Trp Ser Thr Trp Thr Glu Trp Ser Val Cys Asn Ser
 275 280 285

65

Arg Cys Gly Arg Gly Tyr Gln Lys Arg Thr Arg Thr Cys Thr Asn Pro
 290 295 300

- 3 -

Ala Pro Leu Asn Gly Gly Ala Phe Cys Glu Gly Gln Ser Val Gln Lys
305 310 315 320

5 Ile Ala Cys Thr Thr Leu Cys Pro Val Asp Gly Arg Trp Thr Pro Trp
325 330 335

Ser Lys Trp Ser Thr Cys Gly Thr Glu Cys Thr His Trp Arg Arg Arg
340 345 350

10 Glu Cys Thr Ala Pro Ala Pro Lys Asn Gly Gly Lys Asp Cys Asp Gly
355 360 365

Leu Val Leu Gln Ser Lys Asn Cys Thr Asp Gly Leu Cys Met Gln Ser
370 375 380

15 Phe Ile Tyr Pro Ile Ser Thr Glu Gln Arg Thr Gln Asn Glu Tyr Gly
385 390 395 400

20 Phe Ser Ser Ala Pro Asp Ser Asp Asp Val Ala Leu Tyr Val Gly Ile
405 410 415

Val Ile Ala Val Ile Val Cys Leu Ala Ile Ser Val Val Val Ala Leu
420 425 430

25 Phe Val Tyr Arg Lys Asn His Arg Asp Phe Glu Ser Asp Ile Ile Asp
435 440 445

Ser Ser Ala Leu Asn Gly Gly Phe Gln Pro Val Asn Ile Lys Ala Ala
450 455 460

30 Arg Gln Asp Leu Leu Ala Val Pro Pro Asp Leu Thr Ser Ala Ala Ala
465 470 475 480

Met Tyr Arg Gly Pro Val Tyr Ala Leu His Asp Val Ser Asp Lys Ile
485 490 495

35 Pro Met Thr Asn Ser Pro Ile Leu Asp Pro Leu Pro Asn Leu Lys Ile
500 505 510

40 Lys Val Tyr Asn Thr Ser Gly Ala Val Thr Pro Gln Asp Asp Leu Ser
515 520 525

Glu Phe Thr Ser Lys Leu Ser Pro Gln Met Thr Gln Ser Leu Leu Glu
530 535 540

45 Asn Glu Ala Leu Ser Leu Lys Asn Gln Ser Leu Ala Arg Gln Thr Asp
545 550 555 560

Pro Ser Cys Thr Ala Phe Gly Ser Phe Asn Ser Leu Gly Gly His Leu
565 570 575

50 Ile Val Pro Asn Ser Gly Val Ser Leu Leu Ile Pro Ala Gly Ala Ile
580 585 590

55 Pro Gln Gly Arg Val Tyr Glu Met Tyr Val Thr Val His Arg Lys Glu
595 600 605

Thr Met Arg Pro Pro Met Asp Asp Ser Gln Thr Leu Leu Thr Pro Val
610 615 620

60 Val Ser Cys Gly Pro Pro Gly Ala Leu Leu Thr Arg Pro Val Val Leu
625 630 635 640

Thr Met His His Cys Ala Asp Pro Asn Thr Glu Asp Trp Lys Ile Leu
645 650 655

65 Leu Lys Asn Gln Ala Ala Gln Gly Gln Trp Glu Asp Val Val Val Val
660 665 670

- 4 -

Gly Glu Glu Asn Phe Thr Thr Pro Cys Tyr Ile Gln Leu Asp Ala Glu
 675 680 685
 5 Ala Cys His Ile Leu Thr Glu Asn Leu Ser Thr Tyr Ala Leu Val Gly
 690 695 700
 His Ser Thr Thr Lys Ala Ala Ala Lys Arg Leu Lys Leu Ala Ile Phe
 705 710 715 720
 10 Gly Pro Leu Cys Cys Ser Ser Leu Glu Tyr Ser Ile Arg Val Tyr Cys
 725 730 735
 15 Leu Asp Asp Thr Gln Asp Ala Leu Lys Glu Ile Leu His Leu Glu Arg
 740 745 750
 Gln Thr Gly Gly Gln Leu Leu Glu Glu Pro Lys Ala Leu His Phe Lys
 755 760 765
 20 Gly Ser Thr His Asn Leu Arg Leu Ser Ile His Asp Ile Ala His Ser
 770 775 780
 Leu Trp Lys Ser Lys Leu Leu Ala Lys Tyr Gln Glu Ile Pro Phe Tyr
 785 790 795 800
 25 His Val Trp Ser Gly Ser Gln Arg Asn Leu His Cys Thr Phe Thr Leu
 805 810 815
 30 Glu Arg Phe Ser Leu Asn Thr Val Glu Leu Val Cys Lys Leu Cys Val
 820 825 830
 Arg Gln Val Glu Gly Glu Gly Gln Ile Phe Gln Leu Asn Cys Thr Val
 835 840 845
 35 Ser Glu Glu Pro Thr Gly Ile Asp Leu Pro Leu Leu Asp Pro Ala Asn
 850 855 860
 Thr Ile Thr Thr Val Thr Gly Pro Ser Ala Phe Ser Ile Pro Leu Pro
 865 870 875 880
 40 Ile Arg Gln Lys Leu Cys Ser Ser Leu Asp Ala Pro Gln Thr Arg Gly
 885 890 895
 45 His Asp Trp Arg Met Leu Ala His Lys Leu Asn Leu Asp Arg Tyr Leu
 900 905 910
 Asn Tyr Phe Ala Thr Lys Ser Ser Pro Thr Gly Val Ile Leu Asp Leu
 915 920 925
 50 Trp Glu Ala Gln Asn Phe Pro Asp Gly Asn Leu Ser Met Leu Ala Ala
 930 935 940
 Val Leu Glu Glu Met Gly Arg His Glu Thr Val Val Ser Leu Ala Ala
 945 950 955 960
 55 Glu Gly Gln Tyr
 60 <210> 3
 <211> 2780
 <212> DNA
 <213> Homo sapiens
 65 <400> 3
 ctgaaaaacag aagatagaga ggagtctcgg agctcgccat ctccagcgat ctctacattg 60
 ggaaaaaaca tggagtcagc tccggcagcc cccgaccccg ccgcccagga gccaggcagc 120
 agcggcgcgg acgcggccgc cggctccagg gagacccgc tgaaccagga atccgcccgc 180

- 5 -

aagagcgagc cgcctgcccc ggtgcgacaga cagagctatt ccagcaccag cagagatgat 240
 gacttttttc atgaactccc agaaactttt ccttctgac cactgagcc tctgccacat 300
 ttccttattg agcctgaaga agcttatatt gtgaagaata agcctgtgaa cctgtactgt 360
 5 aaagcaagcc ctgccacca gatctatttc aagtgtataa gtgaatgggt tcatcagaag 420
 gaccacatag tagatgaaag agtagatgaa acttccggtc tcattgtccg ggaagtgagc 480
 attgagattt cgcgccagca agtgaagaa ctctttggac ctgaagatta ctggtgccag 540
 tgtgtggcct ggagctccgc ggttaccaca aagagccgga agcggtatgt gcgcattgca 600
 tatctacgga agacatttga gcaggaaccc ctaggaaagg aagtgtcttt ggaacaggaa 660
 gtcttactcc agtgtcgacc acctgaaggg atcccagtg ctgaggtgga atgggtgaaa 720
 10 aatgaagaca taattgatcc cgttgaagat cggaattttt atattactat tgatcacaac 780
 ctcatacaaa agcaggcccg actctctgat actgcaaat acacctgtgt tgccaaaaac 840
 attgttgcca agaggaaaag tacaactgcc actgtcatag tctatgtcaa cgggtggctgg 900
 tccacttgga cggagtgggtc tgtgtgtaac agccgctgtg gacgagggtg tcagaaacgt 960
 15 acaaggactt gtaccaaccc ggcaccactc aatgggggtg ccttctgtga agggcagagt 1020
 gtgcagaaaa tagcctgtac tacgttatgc ccagtggatg gcaggtggac gccatggagc 1080
 aagtgttcta cttgtggaac tgagtgcacc cactggcgca ggaggggatg cacggcgcca 1140
 gccccaaaga atggaggcaa ggactgcgac ggctcgtct tgcaatccaa gaactgcact 1200
 gatgggcttt gcatgcagag ttctatttat cccatttcaa ctgaacagag aaccagaaat 1260
 gaatatggat tttctcttgc tctgtattca gatgatgttg ctctctatgt tgggattgtg 1320
 20 atagcagtga tcgtttgcct ggcgactctt gtagttgttg ccttgtttgt gtatcggaag 1380
 aatcatcgtg actttagtc agatattatt gactcttcgg cactcaatgg gggctttcag 1440
 cctgtgaaca tcaaggcagc aagacaagat ctgctggctg tacccccaga cctcacgtca 1500
 gctgcagcca tgtacagagg acctgtctat gccctgcagt acgtctcaga caaatccca 1560
 atgaccaact ctccaattct ggatccactg cccaacctga aaatcaaatgt gtacaacacc 1620
 25 tcaggtgctg tcacccccc agatgacctc tctgagttta cgtccaagct gtcccctcag 1680
 atgaccaggt cgtgttgga gaataagcc ctcagcctga agaaccagag tctagcaagg 1740
 cagactgac catcctgtac cgcatttggc agcttcaact cactgggagg tcaccttatt 1800
 gttcccaatt caggagtcag cttgtgtatt cccgctgggg ccattcccca agggagagtc 1860
 tacgaaatgt atgtgactgt acacaggaaa gaaactatga ggccacccat ggatgactct 1920
 30 cagacacttt tgaccctgtt ggtgagctgt gggcccccag gagctctgct caccgcccc 1980
 gtcgtctca ctatgcatca ctgctgcagac cccaataccg aggactggaa aatactgctc 2040
 ctgctggag tacagcatcc gagtctactg tctggatgac acccaggatg ccctgaagga 2100
 aattttacat cttgagagac agacgggagg acagctccta gaagaacctg aggtctctca 2160
 ttttaaggc agcaccaca acctgcgcct gtcaattcac gatctgccc attccctctg 2220
 35 gaagagcaaa ttgctggcta aatatcagga aattccattt taccatgttt ggagtggatc 2280
 tcaaagaaac ctgcaactgca ccttcaactct ggaaagattt agcctgaaca cagtggagct 2340
 ggtttgcaaa ctctgtgtgc ggcaggtgga aggagaaggg cagatcttcc agctcaactg 2400
 caccgtgtca gaggaacctg ctggcatcga tttgctgctg ctggatctct cgaacaccat 2460
 caccacggtc acggggccca gtgctttcag catccctctc cctatccggc agaagctctg 2520
 40 tagcagcctg gatgcccccc agacgagagg ccatgactgg aggatgctgg ccataagct 2580
 gaacctggac aggtacttga attactttgc caccaaatcc agcccaactg gcgtaactct 2640
 ggatctttgg gaagcacaga acttccaga tggaaacctg agcatgctgg cagctgtctt 2700
 ggaagaaatg ggaagacatg aaacggtggt gtccttagca gcagaagggc agtattaacc 2760
 accatgctgg aaggggaaat 2780

<210> 4

<211> 669

<212> PRT

50 <213> Homo sapiens

<400> 4

Met Glu Ser Ala Pro Ala Ala Pro Asp Pro Ala Ala Ser Glu Pro Gly
 1 5 10 15

55 Ser Ser Gly Ala Asp Ala Ala Ala Gly Ser Arg Glu Thr Pro Leu Asn
 20 25 30

60 Gln Glu Ser Ala Arg Lys Ser Glu Pro Pro Ala Pro Val Arg Arg Gln
 35 40 45

Ser Tyr Ser Ser Thr Ser Arg Asp Asp Asp Phe Phe His Glu Leu Pro
 50 55 60

65 Glu Thr Phe Pro Ser Asp Pro Pro Glu Pro Leu Pro His Phe Leu Ile
 65 70 75 80

Glu Pro Glu Glu Ala Tyr Ile Val Lys Asn Lys Pro Val Asn Leu Tyr

- 6 -

	85	90	95
	Cys Lys Ala Ser Pro Ala Thr Gln Ile Tyr Phe Lys Cys Asn Ser Glu		
	100	105	110
5	Trp Val His Gln Lys Asp His Ile Val Asp Glu Arg Val Asp Glu Thr		
	115	120	125
10	Ser Gly Leu Ile Val Arg Glu Val Ser Ile Glu Ile Ser Arg Gln Gln		
	130	135	140
	Val Glu Glu Leu Phe Gly Pro Glu Asp Tyr Trp Cys Gln Cys Val Ala		
	145	150	155
15	Trp Ser Ser Ala Gly Thr Thr Lys Ser Arg Lys Ala Tyr Val Arg Ile		
	165	170	175
	Ala Tyr Leu Arg Lys Thr Phe Glu Gln Glu Pro Leu Gly Lys Glu Val		
	180	185	190
20	Ser Leu Glu Gln Glu Val Leu Leu Gln Cys Arg Pro Pro Glu Gly Ile		
	195	200	205
25	Pro Val Ala Glu Val Glu Trp Leu Lys Asn Glu Asp Ile Ile Asp Pro		
	210	215	220
	Val Glu Asp Arg Asn Phe Tyr Ile Thr Ile Asp His Asn Leu Ile Ile		
	225	230	235
30	Lys Gln Ala Arg Leu Ser Asp Thr Ala Asn Tyr Thr Cys Val Ala Lys		
	245	250	255
	Asn Ile Val Ala Lys Arg Lys Ser Thr Thr Ala Thr Val Ile Val Tyr		
	260	265	270
35	Val Asn Gly Gly Trp Ser Thr Trp Thr Glu Trp Ser Val Cys Asn Ser		
	275	280	285
40	Arg Cys Gly Arg Gly Tyr Gln Lys Arg Thr Arg Thr Cys Thr Asn Pro		
	290	295	300
	Ala Pro Leu Asn Gly Gly Ala Phe Cys Glu Gly Gln Ser Val Gln Lys		
	305	310	315
45	Ile Ala Cys Thr Thr Leu Cys Pro Val Asp Gly Arg Trp Thr Pro Trp		
	325	330	335
	Ser Lys Trp Ser Thr Cys Gly Thr Glu Cys Thr His Trp Arg Arg Arg		
	340	345	350
50	Glu Cys Thr Ala Pro Ala Pro Lys Asn Gly Gly Lys Asp Cys Asp Gly		
	355	360	365
55	Leu Val Leu Gln Ser Lys Asn Cys Thr Asp Gly Leu Cys Met Gln Ser		
	370	375	380
	Phe Ile Tyr Pro Ile Ser Thr Glu Gln Arg Thr Gln Asn Glu Tyr Gly		
	385	390	395
60	Phe Ser Ser Ala Pro Asp Ser Asp Asp Val Ala Leu Tyr Val Gly Ile		
	405	410	415
	Val Ile Ala Val Ile Val Cys Leu Ala Ile Ser Val Val Val Ala Leu		
	420	425	430
65	Phe Val Tyr Arg Lys Asn His Arg Asp Phe Glu Ser Asp Ile Ile Asp		
	435	440	445

- 7 -

Ser Ser Ala Leu Asn Gly Gly Phe Gln Pro Val Asn Ile Lys Ala Ala
450 455 460

Arg Gln Asp Leu Leu Ala Val Pro Pro Asp Leu Thr Ser Ala Ala Ala
5 465 470 475 480

Met Tyr Arg Gly Pro Val Tyr Ala Leu His Asp Val Ser Asp Lys Ile
485 490 495

10 Pro Met Thr Asn Ser Pro Ile Leu Asp Pro Leu Pro Asn Leu Lys Ile
500 505 510

Lys Val Tyr Asn Thr Ser Gly Ala Val Thr Pro Gln Asp Asp Leu Ser
15 515 520 525

Glu Phe Thr Ser Lys Leu Ser Pro Gln Met Thr Gln Ser Leu Leu Glu
530 535 540

20 Asn Glu Ala Leu Ser Leu Lys Asn Gln Ser Leu Ala Arg Gln Thr Asp
545 550 555 560

Pro Ser Cys Thr Ala Phe Gly Ser Phe Asn Ser Leu Gly Gly His Leu
565 570 575

25 Ile Val Pro Asn Ser Gly Val Ser Leu Leu Ile Pro Ala Gly Ala Ile
580 585 590

Pro Gln Gly Arg Val Tyr Glu Met Tyr Val Thr Val His Arg Lys Glu
595 600 605

30 Thr Met Arg Pro Pro Met Asp Asp Ser Gln Thr Leu Leu Thr Pro Val
610 615 620

Val Ser Cys Gly Pro Pro Gly Ala Leu Leu Thr Arg Pro Val Val Leu
35 625 630 635 640

Thr Met His His Cys Ala Asp Pro Asn Thr Glu Asp Trp Lys Ile Leu
645 650 655

40 Leu Leu Ala Gly Val Gln His Pro Ser Leu Leu Ser Gly
660 665

<210> 5
45 <211> 1863
<212> DNA
<213> Homo sapiens

<400> 5
50 gttttttttt ttttttttgg accatagagt tgctctgaaa acagaagata gagggagtct 60
cggagctcgc atctccagcg atctctacat tgggaaaaaa catggagtca gctccggcag 120
ccccgaccc cgcgcgccag gagccaggca gcagcggcgc ggacgcggcc gccggctcca 180
gggagacccc gctgaaccag gaatccgccc gcaagagcga gccgcctgcc ccggtgcgca 240
gacagagcta ttccagcacc agcagaggta tctcagtaac gaagaagaca catacatctc 300
55 aaattgaaat tattccatgc aagatctgtg gagacaaatc atcaggaatc cattatggtg 360
tcattacatg tgaaggctgc aagggctttt tcaggagaag tcagcaaagc aatgccacct 420
actcctgtcc tcgtcagaag aactgtttga ttgatcgaac cagtagaaac cgctgccaac 480
actgtcgatt acagaaatgc cttgccgtag ggatgtctcg agatgctgta aaatttggcc 540
gaatgtcaaa aaagcagaga gacagcttgt atgcagaagt acagaaacac cggatgcagc 600
60 agcagcagcg cgaccaccag cagcagcctg gagaggctga gccgctgacg cccacctaca 660
acatctcggc caacgggctg acggaacttc acgacgacct cagtaactac attgacgggc 720
acacccctga ggggagtaag gcagactccg ccgtcagcag cttctacctg gacatacagc 780
cttccccaga ccagtcaggt cttgatatca atggaatcaa accagaacca atatgtgact 840
acacaccagc atcaggcttc ttccctact gtctgttcac caacggcgag acttccccaa 900
65 ctgtgtccat ggcagaatta gaacaccttg cacagaatat atctaaatcg catctggaaa 960
cctggcaata cttgagagaa gagctccagc agataacgtg gcagaccttt ttacaggaag 1020
aaattgagaa ctatcaaaac aagcagcggg aggtgatgtg gcaattgtgt gccatcaaaa 1080
ttacagaagc tatacagtat gtggtggagt ttgccaaacg cattgatgga tttatggaac 1140

- 8 -

5 t g t g t c a a a a t g a t c a a a t t g t g c t t c t a a a a g c a g g t t c t c t a g a g t g g t g t t t t a t c a 1200
 g a a t g t g c c g t g c c t t t g a c t c t c a g a a c a c a c c g t g t a c t t t g a t g g g a a g t a t g c c a 1260
 g c c c c g a c g t c t t c a a t c c t t a g g t t g t a a g a c t t t a t t a g c t t t g t g t t g a a t t t g 1320
 g a a g a g t t t a t g t t c t a t g c a c c t g a c t g a a g a t g a a t t g c a t t a t t t t c t g c a t t t g 1380
 t a c t g a t g t c a g c a g a t c g c t c a t g g c t g c a a g a a a g g t a a a a a t t g a a a a c t g c a a c 1440
 a g a a a a t t c a g c t a g c t c t c a a c a c g t c c t a c a g a a g a t c a c c g a g a a g a t g g a a t a c 1500
 t a a c a a a g t t a a t a t g c a a g g t g t c t a c a t t a a g a c c t a t g t g g a c g a c a t a c a g a a a 1560
 a g c t a a t g g c a t t t a a g c a a t a c c c a g a c a t t g t g c g a c t t c a t t t t c c t c c a t t a t 1620
 a c a a g g a g t t g t t c a c t t c a g a a t t g a g c a g c a a t g c a a a t t g a t g g g t a a a t g t t a t 1680
 10 c a c c t a a g c a c t t c t a g a a t g t c t g a a g t a c a a c a t g a a a a a c a a c a a a a a a t t a a c 1740
 c g a g a c a c t t t a t a t g g c c c t g c a c a g a c c t g g a g c g c c a c a c a c t g c a c a t c t t t t g g t 1800
 g a t c g g g g t c a g g c a a a g g a g g g g a a c a a t g a a a c a a a t a a g t t g a a c t t g t t t t t c 1860
 t c a 1863

15 <210> 6
 <211> 523
 <212> PRT
 <213> Homo sapiens

20 <400> 6
 Met Glu Ser Ala Pro Ala Ala Pro Asp Pro Ala Ala Ser Glu Pro Gly
 1 5 10 15

25 Ser Ser Gly Ala Asp Ala Ala Ala Gly Ser Arg Glu Thr Pro Leu Asn
 20 25 30

Gln Glu Ser Ala Arg Lys Ser Glu Pro Pro Ala Pro Val Arg Arg Gln
 35 40 45

30 Ser Tyr Ser Ser Thr Ser Arg Gly Ile Ser Val Thr Lys Lys Thr His
 50 55 60

Thr Ser Gln Ile Glu Ile Ile Pro Cys Lys Ile Cys Gly Asp Lys Ser
 35 65 70 75 80

Ser Gly Ile His Tyr Gly Val Ile Thr Cys Glu Gly Cys Lys Gly Phe
 85 90 95

40 Phe Arg Arg Ser Gln Gln Ser Asn Ala Thr Tyr Ser Cys Pro Arg Gln
 100 105 110

Lys Asn Cys Leu Ile Asp Arg Thr Ser Arg Asn Arg Cys Gln His Cys
 115 120 125

45 Arg Leu Gln Lys Cys Leu Ala Val Gly Met Ser Arg Asp Ala Val Lys
 130 135 140

Phe Gly Arg Met Ser Lys Lys Gln Arg Asp Ser Leu Tyr Ala Glu Val
 50 145 150 155 160

Gln Lys His Arg Met Gln Gln Gln Gln Arg Asp His Gln Gln Gln Pro
 165 170 175

55 Gly Glu Ala Glu Pro Leu Thr Pro Thr Tyr Asn Ile Ser Ala Asn Gly
 180 185 190

Leu Thr Glu Leu His Asp Asp Leu Ser Asn Tyr Ile Asp Gly His Thr
 195 200 205

60 Pro Glu Gly Ser Lys Ala Asp Ser Ala Val Ser Ser Phe Tyr Leu Asp
 210 215 220

Ile Gln Pro Ser Pro Asp Gln Ser Gly Leu Asp Ile Asn Gly Ile Lys
 65 225 230 235 240

Pro Glu Pro Ile Cys Asp Tyr Thr Pro Ala Ser Gly Phe Phe Pro Tyr
 245 250 255

- 9 -

Cys Ser Phe Thr Asn Gly Glu Thr Ser Pro Thr Val Ser Met Ala Glu
 260 265 270

5 Leu Glu His Leu Ala Gln Asn Ile Ser Lys Ser His Leu Glu Thr Cys
 275 280 285

Gln Tyr Leu Arg Glu Glu Leu Gln Gln Ile Thr Trp Gln Thr Phe Leu
 290 295 300

10 Gln Glu Glu Ile Glu Asn Tyr Gln Asn Lys Gln Arg Glu Val Met Trp
 305 310 315 320

Gln Leu Cys Ala Ile Lys Ile Thr Glu Ala Ile Gln Tyr Val Val Glu
 325 330 335

15 Phe Ala Lys Arg Ile Asp Gly Phe Met Glu Leu Cys Gln Asn Asp Gln
 340 345 350

20 Ile Val Leu Leu Lys Ala Gly Ser Leu Glu Val Val Phe Ile Arg Met
 355 360 365

Cys Arg Ala Phe Asp Ser Gln Asn Asn Thr Val Tyr Phe Asp Gly Lys
 370 375 380

25 Tyr Ala Ser Pro Asp Val Phe Lys Ser Leu Gly Cys Glu Asp Phe Ile
 385 390 395 400

Ser Phe Val Phe Glu Phe Gly Lys Ser Leu Cys Ser Met His Leu Thr
 405 410 415

30 Glu Asp Glu Ile Ala Leu Phe Ser Ala Phe Val Leu Met Ser Ala Asp
 420 425 430

35 Arg Ser Trp Leu Gln Glu Lys Val Lys Ile Glu Lys Leu Gln Gln Lys
 435 440 445

Ile Gln Leu Ala Leu Gln His Val Leu Gln Lys Asn His Arg Glu Asp
 450 455 460

40 Gly Ile Leu Thr Lys Leu Ile Cys Lys Val Ser Thr Leu Arg Ala Leu
 465 470 475 480

Cys Gly Arg His Thr Glu Lys Leu Met Ala Phe Lys Ala Ile Tyr Pro
 485 490 495

45 Asp Ile Val Arg Leu His Phe Pro Pro Leu Tyr Lys Glu Leu Phe Thr
 500 505 510

50 Ser Glu Phe Glu Pro Ala Met Gln Ile Asp Gly
 515 520

<210> 7

55 <211> 1821

<212> DNA

<213> Homo sapiens

<400> 7

60 ctagattcct ctgtgtttca taccaggag gtcaggagg gcaccttgta gtgcctgagg 60
 ccccgagagg gttcacagcc acctgagaag tgtttgcagt ctatgggtgt gtgttcttc 120
 tccaagctgt gtcagctgtg ggaatggcgtg gagagaaaac aagaaaaagt tgttccctc 180
 agctctctgc tgaaaaaatg ctcacgggga gtcagtcaac atttggtatc tcagtaacga 240
 agaagacaca tacatctcaa attgaaatta ttccatgcaa gatctgtgga gacaaatcat 300
 65 caggaatcca ttatgggtgc attacatgtg aaggctgcaa gggctttttc aggagaagtc 360
 agcaaagcaa tgccacctac tcctgtcctc gtcagaagaa ctgtttgatt gatcgaacca 420
 gtagaaaccg ctgccaacac tgcgattac agaaatgcct tgccgtaggg atgtctcgag 480
 atgctgtaaa atttggccga atgtcaaaaa agcagagaga cagcttgat gcagaagtac 540

- 10 -

agaaacaccg gatgcagcag cagcagcgcg accaccagca gcagcctgga gaggtgagc 600
 cgctgacgcc caccctacaac atctcggcca acgggctgac ggaacttcac gacgacctca 660
 gtaactacat tgacggggcac acccctgagg ggagtaaggc agactccgcc gtcagcagct 720
 tctacctgga catacagcct tccccagacc agtcaggctc tgatatcaat ggaatcaaac 780
 5 cagaaccaat atgtgactac acaccagcat caggcttctt tccctactgt tcgttcacca 840
 acggcgagac ttccccaact gtgtccatgg cagaattaga acaccttgca cagaatatat 900
 ctaaactcgca tctggaaacc tgccaatact tgagagaaga gctccagcag ataactgtgc 960
 agaccttttt acaggaagaa attgagaact atcaaaacaa gcagcgggag gtgatgtggc 1020
 aattgtgtgc catcaaaatt acagaagcta tacagtatgt ggtggagttt gccaaacgca 1080
 10 ttgatggatt tatggaactg tgtcaaaatg atcaaatgt gcttctaaaa gcaggttctc 1140
 tagaggtggt gtttatcaga atgtgccgtg cctttgactc tcagaacaac accgtgtact 1200
 ttgatgggaa gtatgccagc cccgacgtct tcaaatcctt aggttggtgaa gactttatta 1260
 gctttgtgtt tgaatttggg aagagtattat gttctatgca cctgactgaa gatgaaattg 1320
 cattattttc tgcatttgta ctgatgtcag cagatcgctc atggctgcaa gaaaaggtaa 1380
 15 aaattgaaaa actgcaacag aaaattcagc tagctcttca acacgtccta cagaagaatc 1440
 accgagaaga ttggaatacta acaaaagtaa tatgcaaggt gtctacatta agagccttat 1500
 gtggacgaca tacagaaaag ctaatggcat ttaaagcaat ataccagac attgtgcgac 1560
 ttcatatttc tccattatac aaggagttgt tcacttcaga atttgagcca gcaatgcaaa 1620
 ttgatgggta aatgttatca cctaagcact tctagaatgt ctgaagtaca aacatgaaaa 1680
 20 acaaacaaaa aaattaaccg agacacttta tatggccctg cacagacctg gagcgccaca 1740
 cactgcacat cttttggtga tcggggtcag gcaaaggagg ggaacaatg aaaacaaata 1800
 aagttgaact tgtttttctc a 1821

25 <210> 8
 <211> 477
 <212> PRT
 <213> Homo sapiens

30 <400> 8
 Met Leu Thr Gly Ser Gln Ser Thr Phe Gly Ile Ser Val Thr Lys Lys
 1 5 10 15
 Thr His Thr Ser Gln Ile Glu Ile Ile Pro Cys Lys Ile Cys Gly Asp
 35 20 25 30
 Lys Ser Ser Gly Ile His Tyr Gly Val Ile Thr Cys Glu Gly Cys Lys
 35 40 45
 40 Gly Phe Phe Arg Arg Ser Gln Gln Ser Asn Ala Thr Tyr Ser Cys Pro
 50 55 60
 Arg Gln Lys Asn Cys Leu Ile Asp Arg Thr Ser Arg Asn Arg Cys Gln
 65 70 75 80
 45 His Cys Arg Leu Gln Lys Cys Leu Ala Val Gly Met Ser Arg Asp Ala
 85 90 95
 Val Lys Phe Gly Arg Met Ser Lys Lys Gln Arg Asp Ser Leu Tyr Ala
 50 100 105 110
 Glu Val Gln Lys His Arg Met Gln Gln Gln Arg Asp His Gln Gln
 115 120 125
 55 Gln Pro Gly Glu Ala Glu Pro Leu Thr Pro Thr Tyr Asn Ile Ser Ala
 130 135 140
 Asn Gly Leu Thr Glu Leu His Asp Asp Leu Ser Asn Tyr Ile Asp Gly
 145 150 155 160
 60 His Thr Pro Glu Gly Ser Lys Ala Asp Ser Ala Val Ser Ser Phe Tyr
 165 170 175
 Leu Asp Ile Gln Pro Ser Pro Asp Gln Ser Gly Leu Asp Ile Asn Gly
 65 180 185 190
 Ile Lys Pro Glu Pro Ile Cys Asp Tyr Thr Pro Ala Ser Gly Phe Phe
 195 200 205

- 11 -

Pro Tyr Cys Ser Phe Thr Asn Gly Glu Thr Ser Pro Thr Val Ser Met
 210 215 220

5 Ala Glu Leu Glu His Leu Ala Gln Asn Ile Ser Lys Ser His Leu Glu
 225 230 235 240

Thr Cys Gln Tyr Leu Arg Glu Glu Leu Gln Gln Ile Thr Trp Gln Thr
 245 250 255

10 Phe Leu Gln Glu Glu Ile Glu Asn Tyr Gln Asn Lys Gln Arg Glu Val
 260 265 270

15 Met Trp Gln Leu Cys Ala Ile Lys Ile Thr Glu Ala Ile Gln Tyr Val
 275 280 285

Val Glu Phe Ala Lys Arg Ile Asp Gly Phe Met Glu Leu Cys Gln Asn
 290 295 300

20 Asp Gln Ile Val Leu Leu Lys Ala Gly Ser Leu Glu Val Val Phe Ile
 305 310 315 320

Arg Met Cys Arg Ala Phe Asp Ser Gln Asn Asn Thr Val Tyr Phe Asp
 325 330 335

25 Gly Lys Tyr Ala Ser Pro Asp Val Phe Lys Ser Leu Gly Cys Glu Asp
 340 345 350

30 Phe Ile Ser Phe Val Phe Glu Phe Gly Lys Ser Leu Cys Ser Met His
 355 360 365

Leu Thr Glu Asp Glu Ile Ala Leu Phe Ser Ala Phe Val Leu Met Ser
 370 375 380

35 Ala Asp Arg Ser Trp Leu Gln Glu Lys Val Lys Ile Glu Lys Leu Gln
 385 390 395 400

Gln Lys Ile Gln Leu Ala Leu Gln His Val Leu Gln Lys Asn His Arg
 405 410 415

40 Glu Asp Gly Ile Leu Thr Lys Leu Ile Cys Lys Val Ser Thr Leu Arg
 420 425 430

45 Ala Leu Cys Gly Arg His Thr Glu Lys Leu Met Ala Phe Lys Ala Ile
 435 440 445

Tyr Pro Asp Ile Val Arg Leu His Phe Pro Pro Leu Tyr Lys Glu Leu
 450 455 460

50 Phe Thr Ser Glu Phe Glu Pro Ala Met Gln Ile Asp Gly
 465 470 475

<210> 9
 55 <211> 3646
 <212> DNA
 <213> Homo sapiens

<400> 9
 60 ctgcctttgg agaaagtggg gtgtggcgct tggttgtcgt tatttcttcg gactgcttcg 60
 cgggtgcacgg attcagcttc tgcccagtg ggctttcagc tgtttgcgcg tctctctgtc 120
 cccctccct ccccccggca cacctctgtc tacgatgagg aaaggtctgc gggcgacagc 180
 ggcccgctgc ggactgggac tgggatactt gctgcaaatg ctctgtctac ctgccctggc 240
 cctgctcagc gccagcggca ctggctccgc cgccaagat gatgactttt ttcataaact 300
 65 cccagaaact tttcttctg atccacctga gcctctgcca catttctta ttgagcctga 360
 agaagcttat atttgaaga ataagcctgt gaacctgtac tgtaaagcaa gccctgccac 420
 ccagatctat ttcaagtgt atagtgaatg ggttcacag aaggaccaca tagtagatga 480
 aagagtagat gaaacttccg gtctcattgt ccgggaagtg agcattgaga tttegcgcca 540

- 12 -

gcaagtggaa gaactctttg gacctgaaga ttactggtgc cagtgtgtgg cctggagctc 600
 cgcgggtacc acaaagagcc ggaaggcgta tgtgcgcatt gcatactctac ggaagacatt 660
 tgagcaggaa cccttaggaa aggaagtgtc tttggaacag gaagctcttac tccagtgtcg 720
 accacctgaa gggatcccag tggctgaggt ggaatggtt aaaaatgaag acataattga 780
 5 tcccgttgaa gatcgggaatt tttatattac tattgatcac aacctcatca taaagcaggc 840
 ccgactctct gatactgcaa attacacctg tgttgccaaa aacattgttg ccaagaggaa 900
 aagtacaact gccactgtca tagtctatgt caacgggtggc tgggtccacct ggacggagtg 960
 gtctgtgtgt aacagccgct gtggacgagg gtatcagaaa cgtacaagga cttgtacca 1020
 cccggcacca ctcaatggcg gtgccttctg tgaaggcgag agtgtgcaga aaatagcctg 1080
 10 tactacgtta tgcccagtg atggcagggt gacgccatgg agcaagtggg ctacttgtgg 1140
 aactgagtgcc acccactggc gcaggaggga gtgcacggcg ccagcccccaga aatggagg 1200
 caaggactgc gacggcctcg tcttgcaatc caagaactgc actgatgggc tttgcatgca 1260
 gactgtctct gattcagatg atgtgtctct ctatgttggg attgtgatag cagtgtatcg 1320
 ttgctggcg atctctgtag ttgtggcctt gtttgtgtat cgggaagaatc atcgtgactt 1380
 15 tgagtcatag attattgact cttcggcact caatggggcg tttcarctg tgaacatcaa 1440
 ggcagcaaga caagatctgc tggctgtacc cccagacctc acgtcagctg cagccatgta 1500
 cagaggacct gtctatggcg tgcacacgtc ctacagacaa atcccaatga ccaactctcc 1560
 aattctggat ccactgcccc acctgaaaat caaagtgtac aacacctcag gtgctgtctc 1620
 cccccaagat gacctctctg agtttacgtc caagctgtcc cctcagatga cccagtcgtt 1680
 20 gttggagaat gaagccctca gcctgaagaa ccagagtcta gcaaggcaga ctgatccatc 1740
 ctgtaccgca tttggcagc tcaactcgct gggagggtcac cttattgttc ccaattcagg 1800
 agtcagcttg ctgattcccc ctggggccat tccccaaggg agagtctacg aatgtatgt 1860
 gactgtacac aggaagaaaa ctatgaggcc acccatggat gactctcaga cacttttgac 1920
 ccctgtgggt agctgtgggc cccagaggagc tctgtctacc cgtcccgctg tccctactat 1980
 25 gcactactgc gcagacccca ataccgagga ctggaaaata ctgtcaaga accaggcagc 2040
 acagggacag tgggaggatg tgggtgggt cggggaggaa aacttcacca cccctgcta 2100
 cattaagctg gatgcagagg cctgccacat cctcacagag aacctcagca cctacgcct 2160
 ggttaggacat tccaccacca aagcggctgc aaagcgcctc aagctggcca tctttggg 2220
 cctgtgtctg tctctgtctg agtacagcat ccgagtctac tgtctggatg acaccagga 2280
 30 cgccttgaag gaaattttac atcttgagag acagacggga ggacagctcc tagaagaacc 2340
 taaggctctt cattttaaag gcagacccca caacctgcgc ctgtcaatc acgatatcgc 2400
 ccattccctc tgggaagagca aattgtctgc taaatatcag gaaattccat tttaccatgt 2460
 ttggagtggg tctcaaagaa acctgcactg cactttcact ctggaaagat ttagcctgaa 2520
 cacagtggag ctgggtttgca aactctgtgt gcggcagggt gaaggagaag ggcagatctt 2580
 35 ccagctcaac tgcaccgtgt cagaggaacc tactggcatc gatttgccgc tgcgtggatc 2640
 tgcgaacacc atcaccacgg tcacggggcc cagtgtcttc agcatccctc tccctatccg 2700
 gcagaagctc tgtagcagcc tggatgccc ccagacgaga ggccatgact ggaggatgct 2760
 ggcccataag ctgaaacttg acaggtactt gaattacttt gccaccaa atccagcccaac 2820
 tggcgtaatc ctggatcttt ggggaagcaca gaacttccca gatggaaacc tgagcatgct 2880
 40 ggcagctgtc ttggaagaaa tgggaagaca tgaacgggtg gtgtccttag cagcagaagg 2940
 gcagtattaa ccaccatgtc ggaaggggaa atgaaggaca aaaatgcaca gggagtctgt 3000
 ggccgtccag gtgaatcaca gctgaggagg aaatccagat gagaccaatg cacttcacag 3060
 gcaagaatgc agcaggagcc agaaggaaaa cagatacaac tgcccattga catgcccact 3120
 ttactcggag atcatcacgg gagttaagaa aaattgtgta aatttgtacc ttgaatttag 3180
 45 ctatcaacct aattttctc ttagtgtggc tgtatgtgt gtggtacagg atcttacagt 3240
 ttcctaggaa acgtttttta ttgctatcca gatatatgga taaactttct taacaaacc 3300
 aattttctaca aatgttgttt acatcaaatt ggacagggat gcagacactg tccatggctc 3360
 gttctatttt tgttcaaatc atttgaagtt gaagctgtgg acggtttgtt gtgtctat 3420
 cagattagta atttacagag aaatcacaga cttttgctac aaatcgtgtg catcaagtgt 3480
 50 ctcagataat cctcccatca gtgttctgtt tctagaacct gtagaaccag tgttactgtt 3540
 tgtatcaggg aagtggagaa tctaagtgt aaaaagaaat aactaagact cctattcctt 3600
 ggagggaccc tctgggtgcc ctttgggaat aaagctgtag cactgc 3646

55 <210> 10
 <211> 931
 <212> PRT
 <213> Homo sapiens

60 <400> 10
 Met Arg Lys Gly Leu Arg Ala Thr Ala Ala Arg Cys Gly Leu Gly Leu
 1 5 10 15

65 Gly Tyr Leu Leu Gln Met Leu Val Leu Pro Ala Leu Ala Leu Leu Ser
 20 25 30

Ala Ser Gly Thr Gly Ser Ala Ala Gln Asp Asp Asp Phe Phe His Glu
 35 40 45

- 13 -

Leu Pro Glu Thr Phe Pro Ser Asp Pro Pro Glu Pro Leu Pro His Phe
 50 55 60

5 Leu Ile Glu Pro Glu Glu Ala Tyr Ile Val Lys Asn Lys Pro Val Asn
 65 70 75 80

Leu Tyr Cys Lys Ala Ser Pro Ala Thr Gln Ile Tyr Phe Lys Cys Asn
 85 90 95

10 Ser Glu Trp Val His Gln Lys Asp His Ile Val Asp Glu Arg Val Asp
 100 105 110

15 Glu Thr Ser Gly Leu Ile Val Arg Glu Val Ser Ile Glu Ile Ser Arg
 115 120 125

Gln Gln Val Glu Glu Leu Phe Gly Pro Glu Asp Tyr Trp Cys Gln Cys
 130 135 140

20 Val Ala Trp Ser Ser Ala Gly Thr Thr Lys Ser Arg Lys Ala Tyr Val
 145 150 155 160

Arg Ile Ala Tyr Leu Arg Lys Thr Phe Glu Gln Glu Pro Leu Gly Lys
 165 170 175

25 Glu Val Ser Leu Glu Gln Glu Val Leu Leu Gln Cys Arg Pro Pro Glu
 180 185 190

30 Gly Ile Pro Val Ala Glu Val Glu Trp Leu Lys Asn Glu Asp Ile Ile
 195 200 205

Asp Pro Val Glu Asp Arg Asn Phe Tyr Ile Thr Ile Asp His Asn Leu
 210 215 220

35 Ile Ile Lys Gln Ala Arg Leu Ser Asp Thr Ala Asn Tyr Thr Cys Val
 225 230 235 240

Ala Lys Asn Ile Val Ala Lys Arg Lys Ser Thr Thr Ala Thr Val Ile
 245 250 255

40 Val Tyr Val Asn Gly Gly Trp Ser Thr Trp Thr Glu Trp Ser Val Cys
 260 265 270

45 Asn Ser Arg Cys Gly Arg Gly Tyr Gln Lys Arg Thr Arg Thr Cys Thr
 275 280 285

Asn Pro Ala Pro Leu Asn Gly Gly Ala Phe Cys Glu Gly Gln Ser Val
 290 295 300

50 Gln Lys Ile Ala Cys Thr Thr Leu Cys Pro Val Asp Gly Arg Trp Thr
 305 310 315 320

Pro Trp Ser Lys Trp Ser Thr Cys Gly Thr Glu Cys Thr His Trp Arg
 325 330 335

55 Arg Arg Glu Cys Thr Ala Pro Ala Pro Lys Asn Gly Gly Lys Asp Cys
 340 345 350

60 Asp Gly Leu Val Leu Gln Ser Lys Asn Cys Thr Asp Gly Leu Cys Met
 355 360 365

Gln Thr Ala Pro Asp Ser Asp Asp Val Ala Leu Tyr Val Gly Ile Val
 370 375 380

65 Ile Ala Val Ile Val Cys Leu Ala Ile Ser Val Val Val Ala Leu Phe
 385 390 395 400

Val Tyr Arg Lys Asn His Arg Asp Phe Glu Ser Asp Ile Ile Asp Ser

- 14 -

	405		410		415
	Ser Ala Leu Asn Gly Gly Phe Gln Pro Val Asn Ile Lys Ala Ala Arg				
	420		425		430
5	Gln Asp Leu Leu Ala Val Pro Pro Asp Leu Thr Ser Ala Ala Ala Met				
	435		440		445
10	Tyr Arg Gly Pro Val Tyr Ala Leu His Asp Val Ser Asp Lys Ile Pro				
	450		455		460
	Met Thr Asn Ser Pro Ile Leu Asp Pro Leu Pro Asn Leu Lys Ile Lys				
	465		470		475
15	Val Tyr Asn Thr Ser Gly Ala Val Ser Pro Gln Asp Asp Leu Ser Glu				
	485		490		495
	Phe Thr Ser Lys Leu Ser Pro Gln Met Thr Gln Ser Leu Leu Glu Asn				
	500		505		510
20	Glu Ala Leu Ser Leu Lys Asn Gln Ser Leu Ala Arg Gln Thr Asp Pro				
	515		520		525
25	Ser Cys Thr Ala Phe Gly Ser Phe Asn Ser Leu Gly Gly His Leu Ile				
	530		535		540
	Val Pro Asn Ser Gly Val Ser Leu Leu Ile Pro Ala Gly Ala Ile Pro				
	545		550		555
30	Gln Gly Arg Val Tyr Glu Met Tyr Val Thr Val His Arg Lys Glu Thr				
	565		570		575
	Met Arg Pro Pro Met Asp Asp Ser Gln Thr Leu Leu Thr Pro Val Val				
	580		585		590
35	Ser Cys Gly Pro Pro Gly Ala Leu Leu Thr Arg Pro Val Val Leu Thr				
	595		600		605
40	Met His His Cys Ala Asp Pro Asn Thr Glu Asp Trp Lys Ile Leu Leu				
	610		615		620
	Lys Asn Gln Ala Ala Gln Gly Gln Trp Glu Asp Val Val Val Val Gly				
	625		630		635
45	Glu Glu Asn Phe Thr Thr Pro Cys Tyr Ile Lys Leu Asp Ala Glu Ala				
	645		650		655
	Cys His Ile Leu Thr Glu Asn Leu Ser Thr Tyr Ala Leu Val Gly His				
	660		665		670
50	Ser Thr Thr Lys Ala Ala Ala Lys Arg Leu Lys Leu Ala Ile Phe Gly				
	675		680		685
55	Pro Leu Cys Cys Ser Ser Leu Glu Tyr Ser Ile Arg Val Tyr Cys Leu				
	690		695		700
	Asp Asp Thr Gln Asp Ala Leu Lys Glu Ile Leu His Leu Glu Arg Gln				
	705		710		715
60	Thr Gly Gly Gln Leu Leu Glu Glu Pro Lys Ala Leu His Phe Lys Gly				
	725		730		735
	Ser Thr His Asn Leu Arg Leu Ser Ile His Asp Ile Ala His Ser Leu				
	740		745		750
65	Trp Lys Ser Lys Leu Leu Ala Lys Tyr Gln Glu Ile Pro Phe Tyr His				
	755		760		765

- 15 -

Val Trp Ser Gly Ser Gln Arg Asn Leu His Cys Thr Phe Thr Leu Glu
770 775 780

5 Arg Phe Ser Leu Asn Thr Val Glu Leu Val Cys Lys Leu Cys Val Arg
785 790 795 800

Gln Val Glu Gly Glu Gly Gln Ile Phe Gln Leu Asn Cys Thr Val Ser
805 810 815

10 Glu Glu Pro Thr Gly Ile Asp Leu Pro Leu Leu Asp Pro Ala Asn Thr
820 825 830

Ile Thr Thr Val Thr Gly Pro Ser Ala Phe Ser Ile Pro Leu Pro Ile
835 840 845

15 Arg Gln Lys Leu Cys Ser Ser Leu Asp Ala Pro Gln Thr Arg Gly His
850 855 860

20 Asp Trp Arg Met Leu Ala His Lys Leu Asn Leu Asp Arg Tyr Leu Asn
865 870 875 880

Tyr Phe Ala Thr Lys Ser Ser Pro Thr Gly Val Ile Leu Asp Leu Trp
885 890 895

25 Glu Ala Gln Asn Phe Pro Asp Gly Asn Leu Ser Met Leu Ala Ala Val
900 905 910

Leu Glu Glu Met Gly Arg His Glu Thr Val Val Ser Leu Ala Ala Glu
915 920 925

30 Gly Gln Tyr
930

35 <210> 11
<211> 6277
<212> DNA
<213> Homo sapiens

40 <400> 11
ataaagcaca ttttagcattg ttgtgtttta ataactgcaa atgttctaaa tgtgaataag 60
agcatataat taagattttt tgaataaact cctaaaattg aaatgtttgt gttgtcttta 120
tatcattacc ctaatgccca tttttggttg aggtatttga cagccagaat aatggctccc 180
caaggatgtc catgggtttt gaatttaaga atatgttaca ttacatggca aatctgtaga 240
45 aattaaagta gcagatggat ttaaagtgtc taatctgctg actttaagag aggtagattt 300
tcttggatta ttccattggg tccaatatta tcacaataat ttctaaatgt ggaggaaatt 360
agaagaagtg agtgcattga tacgagaagt acttgaacca tacttgctgg gtgtgaagat 420
gggggaaggg gccactagcc aaagaatgta ggtgtcctct ggaagctgga aacagctaga 480
aaacatgctt tctgtagagg ctgcagaaag gagctcagct ctgggatcat tctcagttcc 540
50 ttaaggctgc cctctggtcc tagccttggc tagcctaggc cttggccttc tcatagaatc 600
acaatttatt ttttaaaaag tcatcaagag aattaatctc tcaagcttgc catgatcatg 660
tcttatgtaa tctaatacaag gaaatggaat atcccccat atccataagc cctggccaca 720
ctcaagagaa aaacatatat atggcaaata tagcagagaa caagaaattt ggaggccatc 780
ttggaattct ctctacaaca ctctctgtat tccctctgtg tgtttgtctt tgtctctctc 840
55 cacatatata tacgtacata cacatccatg cacatatgta tttgtattac ttttttcaat 900
gtagacatat cagcttctta taatatttga tttctttggt tttccactag ttcagtcata 960
tttcgctggt tttagctacca gatccccatc gatacatgaa aacaacatga agaagataaa 1020
gaaagtctct ctatgcattc tttctctcgt gacattcaat tgcatactca ccatcaccca 1080
aacagcaaaa ttattgactg acacttcaaa taatctcttc taatacttat ttacattaaa 1140
60 aatatttttc tttatgaggg tgaagtagaa attaattcac agctctgtac tctggacaaa 1200
gaaaaaactg taacaaaaaa atcttcacac tgatctaaca ctttatattc tcatgagtca 1260
agtgattaat taaattccaa tgctcctctg aaggaggttg ggcatttcta taaattaaaa 1320
aatatcgccg ggcattggtg ctacagctctg taatcccagc actttgggaa gccaaaggcag 1380
gcagatcacg aggtcaggag atcgagacca tcctggctta caggtgaaa ccctgtctct 1440
65 actaaaaata caaaaaaaa aaagtttagg tgggcgtagt ggcgggcgcc tggtagtcca 1500
gctactcggg aggtcgaggc agaagaatgg cgtgaaccca ggaggcggag cttgcagtga 1560
gccaaagatc tgccactgca ctccagcctg ggtgacagag cgagactcca tctcaaaaaa 1620
aaaaaaaaa ttataactaa agttttgaat taaatcatc tgattgtcca ggaatgcaca 1680

cccttaaaata tatgtaataa aaaattaaag atttcttata gaaaaactta ttgaaaaagt 1740
 ccctgaaaga taacaaatga tattataaac tatgtatttc tgtagtaaat ttcccaaaaa 1800
 tacataatatt tatgagcaaa atggtgtgga atatataaaa tattgtatgt atgtgtgtgt 1860
 atataagctt ttttaagtga gtttaagaaac atgaatgagg acctatgaag cattatttct 1920
 5 gtgcacgtac caatactctg ttatttccag tgcccaaccc taagatctct gactctacca 1980
 cattaacttc agactaaaaa ctaacccttt taaacacacc gtaaatgatt tttttttgtt 2040
 aaatgcaact taaaaatttc aaaatatata atagtataaa tgatgtttaa tatattgtgt 2100
 tgaatgttta aagaataaac tgtagtttga atgcatctga atattttgtta ggcatgcaac 2160
 agctcttagt aagaacccat tccattcttc tgagagaaag ttgtatatta actgaagaga 2220
 10 ggcatcttgt tttgatctgg aggacacaaga aatcttaaaag tgtgaatccc attaggttat 2280
 aagataaaaa aaaatcctat cagagaatag gaaaataaaa accactagta atctcattac 2340
 tgcctcttca aatactactg ctatataggt gaacagcttc cctcttagtc tttcctcttt 2400
 ctcccaaaa gtgacgacga ttggtatcaag catgtctagt gggatataag cacaagtat 2460
 gcttattctt gttttctcct ctttttattt gttctgatta ttaatttttc ccctagtcat 2520
 15 agttcttttt ttggatgaaa ttttagggcaa gtgttcaggc ttacctagt acttcagatg 2580
 atgacatctc ctctaaacag taattaatat tttcagtcga tggtaatatt caacatttta 2640
 agctcattta aataactaac tctctgtctc gatattggtt gactctgtgt ctacacccaa 2700
 atctcatgtc aaattgtaat ccccacttat tgaagggtgg gactgggtgg aggtgattgg 2760
 atcatagtcg ggggttccag tgggttagcg ccatcctccc tagggctgtc ttgtgataga 2820
 20 gttctcacca gatccgatga gatctgatgg tttcaaagtg tgtagcagtc tcagtctctc 2880
 tgtctgtctg tctgtctctc tctctgtctc tcaatccct gctccacat ggtaagatgt 2940
 gcttacttcc cctttatgtt ttgacacgat tgtaagtttc ctgaggctc ccagccatgc 3000
 tctgtacag cctgaggaaac tgtgagtcga ttaaatgtct tttcttcata aactaccag 3060
 tctcagtag ttctttatag cagcatgaaa acagactaat acacccctt tctcagcctg 3120
 25 tgcttgttgt aaggctagcc aatgcaagca gaaagagggg tgctgtctgc tgtttctcca 3180
 attggcatct gtgccccttc ttcatttgtt cagttatagt gggaggccca agacttactt 3240
 ttggcctcag acacagtcga ggatcatcct gatgcaaaaa acaaaatcac ctttgaacct 3300
 ctctctccac accatcttcc acttttggtt tccgctatgc ctttgtactg gggacaaaga 3360
 ctatccacga atccttactt taaaaactga ggcagtggct ctgaatcatt ccccttgga 3420
 30 cacccccaca catctgaaca cctcgtgtc ccaacaccac atctgctgaa gttgactgga 3480
 cctgaaatgc tggaggctca gtgttttcgg aaactgcgga gagttctcct gccacaattc 3540
 tcttgtgtgc cgtcaaaagc aattacatgt ttcagatgat gaaaagcact gtcataaaaa 3600
 gctggcacag cagccaagga gacagatgta ttaaacacga cggcttctca gaactcccac 3660
 ttacgttgga gctcagcaca gagatgagca tgggtatata ccccgctcta acatcctccc 3720
 35 atttggattt aatgcatgag gctcgtttct gatgaggtgc caaattcttt cgggtggaca 3780
 acacaaagt gagaatttca gtgatcatac aagtccatgt cctcaaacga aaataataat 3840
 actataaatg caaagatgt accataaaaa atacagtagg tggacttttg tgagtcaaca 3900
 taaaaaaatc cagctattct aactggtagc atgaggagta ccagtctgca ggatgttaac 3960
 gcatgatag gctatggaca atatatcca cacttacgga agtggggaaa ctataggta 4020
 40 gacaaagcta cagggtactt tttctttctt ttttcttttt tttctttctga gatggagtct 4080
 cctctgtctg cctaggctgg agtgcaatgg catgatctca gctcactgca acctccgct 4140
 cccgggttca agtgattctc ctgctcagc ctctgagtc gctgggatta taggcgcaca 4200
 ccaccaacca cccggctaatt ttttcatatt ttttagtagaa acgaggtttc actgtgttag 4260
 ccagaatggt ctgcaactcc tgacctaggt gatctgctc cctcagcctc ccaaagtgt 4320
 45 gggattacag gtgtgagcca cggcactcgg cccaagtact tcttaactgc aggtacttct 4380
 agaggctttt aatatgttaa tgagtgtgt gacgtcagat gactcaaatc aatatatagc 4440
 caacacgttt gatatgacca cagaacctat ttcctgaac actttttatc acctcctaga 4500
 gcaggggtca gaaaactact gccatgtgg tctgatctgg ccactgcttg ttttcgtaaa 4560
 ttttcaactg ccacagccaa cctcatttgt gtacatgccc tccatggctg ctttcacact 4620
 50 gcagccgcag cagtgaagca tgagacagc acccggtggt cacaagatct aaaatattta 4680
 ctctttggcc ctttacggaa aacgtgtaca aactttgtcc tctgaaatat agtttgggaa 4740
 cctacttcac tcatgatata aataattcta aaaaatttca aacaattttt aaagtcaaca 4800
 ttgagaagtt ttccatggtt agaatttaat aattacagtc aataactctc ttaagcactc 4860
 gtgacaataa agtagaagga cactgtgtta tcttatatgg tctctgtctt caggagctc 4920
 55 ccaagctagt ttgtgaaaaa gagaactcaa ggcacaaaca actggggggc aagagatgtt 4980
 ctaaacacta agaggaggca atatggcata gtgaaagaa tgcaagcatg tgtaagagta 5040
 aaacatacac agggcgacga gtcacaagag tcaggggttg gtcttggttg tgatacctc 5100
 tcctgcgata agtagtaagc ctgctttatt gagcacttac tatgtgccag gcaccatccc 5160
 agttgctttc catgaaggat tgctctttaa ttttatcat aatgaatagg tactattatt 5220
 60 atccagttt atgggtgagg aaaccaagtc ttagggtgac taacttatta aaataaagca 5280
 catacctagt aggtggtgaa caagagccaa aacaagctg ttggtcagct aaccaccag 5340
 ttatagttgc tgcgtcttag gggcagcata tctcctctgc atttccccag aggtgacaga 5400
 agtagtgagg aggttagtgc tctctaggct ctgacacaca gccatgttt gctgactcaa 5460
 caaccacaca acgtaatgct atggcacctc aggtggcagg gagggtatcg tgggctggac 5520
 65 tgatcttcag aaggctgcat ctgagcaggg tgtggaagga tggcagaagg aagaggggag 5580
 tgaggcactc aggttagaca gaaggtgata ctgccacatc cccactcag aaggaactc 5640
 gaaggttttg acaaatatgt acatttttga gtgggaacag cagaaagtag gggctggcta 5700
 cctactttcc actaccctta tccccactcc ccttaggggg tgaggcagag ataagcacgc 5760

- 17 -

5 tttgtggggc acactgatcc acaaacttgc aagcctgagc tccactcagc agcagcctgg 5820
 gcaggatgaa tgggtggttg cagggtgttg cggcatgcat agcttctctt tctcctagag 5880
 tcaatccagt gatccccctgc gaaaagaatc ctccacttga cctgaatgct ctgtattggc 5940
 accttctgga caggaatcct atcgggtatt tcatcaaagc atcccacagc agttttaagt 6000
 gtggctccat gagagtaaga caccaggaca gctgtgccac ctagtgatgc ttgcctgtac 6060
 tgtctgcatc ccagagagtg ggagaggctt cggaaaggac tggattttgc tatttttggc 6120
 caacacttct acatgcaaag gaatctgaag tctccgcctg ggcttcaaaa catttcttgt 6180
 ggcctcactg accattcccg tgaacctcag gggacatgag tgtcttcaag ctaagaggga 6240
 gcagtccaac tctgcagcac gaattaccaa ggaacat 6277

10

<210> 12
 <211> 1829
 <212> DNA

15 <213> Homo sapiens

<400> 12
 20 ttttgccttg gttttccag ctattaaaag ggagatatca tcattcattc aacaagcatt 60
 tcttgccggc ccattttcca cgacagaacc atccaaacca agacacttct tctatttgtg 120
 gcaaaaatga gcaacgttaa aactccccag aatgcagagg tgactcacgt gtttatgtaa 180
 gtaagcattt agatatggct ttggaaaaga aaacacacac ctaaaagagg ctaccagcat 240
 ttgaataaag tgtgtaacct ggggtttcaa aactgtacaa cagtttctgt gtactttaag 300
 tcactatcct tcatacttct tccctttaaa ttctatttga aactaaataa tatcaacaat 360
 actttaagta agattatgct ggtgttactg atgaaattat atgcaggctc gcttggcctg 420
 25 tcttcctgcy taagcctttg ataacttgtt cctaggttat ccagttctgc ttacctgagt 480
 acccacttta tcccagaaac gagacctgaa cactgtgtcc tggtttcatc tttccagtga 540
 gccagccttc tgaggcatta gcggctcagc aagaccatgg aggtgagaga ggcacaggca 600
 cactccagat gcagggtctc tgccctctgt caagagtcac acggcagaga ggacaagact 660
 cctgtcccgg gggaactgtt tttccttttg cccacacccc cagatatott cgtgacagcc 720
 30 cctgccctga aagcccggtg agcgtcatgg cggagaagcc agcctgtttc cgattattct 780
 gcctgacagg gactatctgc tgctgtgtgc agttgtagaa atgtcataag ctccaggactc 840
 aggacacttt gcagtcgagt ccaacaggac ccatagagtc cccctgaaca atttctcatc 900
 gtgtgggaag cctaagggtt ggttttaagt aggggctggc agggagtgga ggtctgggct 960
 gcttctttac ggtctttcaa gtgcccggca tttcctggga ttcccaggag taagtcttga 1020
 35 ctcccatgaa tcccactgac ccctaagtggt catgggaaag gagaggacaa aaagaaagga 1080
 gttagacatt tgactaatac tgttcagagt tgaatgcgc tcttttagca ggtaggcatt 1140
 taaagctaac ttttctctac ttgtctactc tgtgggttct ttggagactg atctgcgtga 1200
 ccccaatggc tcaacagaaa tcttctccta caattccctt gactaggtaa tgccctcact 1260
 ctcttgctat cttcttctcc ctggttctca gtgtttcact gccacataaa ctctctctct 1320
 40 cactgtttgg gtcccatgac ccctctgagt ggaattctct ggcaaaagat tcaatgcaag 1380
 ctccaccagg tatctcacat tgggcttaca tctggtctcc tggaaatagc cccatctct 1440
 ttgtctggac aatctgaga cagaccagc ccaatgcaga agtgcttatt cttctctcag 1500
 cagcagatta acttgtctgg gaggtaactc tctcagact tctcaagctc attccgacac 1560
 ctgtccacct aatgcagcaa ctgcacgtga cacaataaac acttaatgca agtcacaaat 1620
 45 gggagtcaca tatataatat tattagtagc cagcctagaa aatgagaca gaaacagata 1680
 aagttaattt caatttcagt aacatttcat ttaattctaat ctgtcaaaaa tactattgtc 1740
 tcaatatata atcaacaaaa ctgttaatga aatacattac agtttttttg tattaagtct 1800
 ttgaaatgga gtgtgtattt tagacttaa 1829

50

<210> 13
 <211> 2612
 <212> DNA
 <213> Homo sapiens

55

<400> 13
 60 atgaggaaag gtctgccggc gacagcggcc cgctgccggc tgggactggg atacttgctg 60
 caaatgctcg tgctacctgc cctggccctg ctccagcgcca ggggactggg ctccgcccgc 120
 caagatgatg acttttttca tgaactccca gaaacttttc cttctgatcc acctgagcct 180
 ctgccacatt tccttattga gcctgaagaa gcttatattg tgaagaataa gcctgtgaac 240
 ctgtactgta aagcaagccc tgccaccag atctatttca agtgtaatag tgaatgggtt 300
 catcagaagg accacatagt agatgaaaga gtagatgaaa cttccgggtc cattgtccgg 360
 gaagttagca ttgagatttc gcgcagcaa gtggaagaac tctttggacc tgaagattac 420
 tgggtgccagt gtgtggcctg gagctccgcy ggtaccacaa agagccggaa ggcgtatgtg 480
 65 cgcattgcat atctacggaa gacatttgag caggaacccc taggaaagga agtgctcttg 540
 gaacagggaag tcttactcca gtgtcgacca cctgaaggga tcccagtgcc tgaggtggaa 600
 tgggtgaaaa atgaagacat aattgatccc gttgaagatc ggaattttta tattactatt 660
 gatcacaacc tcatcataaa gcaggcccca ctctctgata ctgcaaatca cacctgtgtt 720

- 18 -

gccaaaaaca ttgttgccaa gaggaagaagt acaactgcca ctgtcatagt ctatgtcaac 780
 ggtggctggt ccacctggac ggagtgggtct gtgtgtaaca gccgctgtgg acgaggggat 840
 cagaaacgta caaggacttg taccaacccg gcaccactca atgggggtgc cttctgtgaa 900
 5 gggcagagtg tgcagaaaat agcctgtact acgttatgcc cagtggatgg caggtggacg 960
 ccattggagca agtggctctac ttgtggaact gactgcaccc actggcgag gagggagtgc 1020
 acggcgccag cccccaagaa tggaggcaag gactgcgacg gcctcgtctt gcaatccaag 1080
 aactgcactg atgggctttg catgcagact gctcctgatt cagatgatgt tgctctctat 1140
 gttgggattg tgatagcagt gatcgtttgc ctggcgatct ctgtagttgt ggccttgttt 1200
 gtgtatcgga agaattcatc tgactttgag tcagatatta ttgactcttc ggcactcaat 1260
 10 ggggggtttc agcctgtgaa catcaaggca gcaagacaag atctgctggc tgtaccccca 1320
 gacctcacgt cagctgcagc catgtacaga ggacctgtct atgccctgca tgactctca 1380
 gacaaaatcc caatgaccaa ctctccaatt ctggatccac tgcccaacct gaaaatcaaa 1440
 gtgtacaaca cctcaggtgc tgtcaccccc caagatgacc tctctgagtt tacgtccaag 1500
 ctgtccctc agatgaccaa gtcgttgttg gagaatgaag ccctcagcct gaagaaccag 1560
 15 agtctagcaa ggcagactga tccatcctgt accgcatttg gcagcttcaa ctactggga 1620
 ggtcacctta ttgttcccaa ttcaggagtc agcttgctga ttcccgctgg ggcattccc 1680
 caaggagag tctacgaat gtatgtgact gtacacagga aagaactat gaggccacc 1740
 atggatgact ctcagacact tttgacctt gtgtgagct gtgggcccc aggagctctg 1800
 ctacccgcc ccgtcgtct cactatgcat cactgcgag accccaatac cgaggactgg 1860
 20 aaaaactgct tctcgtctg agtacagcat ccgagtctac tgtctggatg acaccagga 1920
 tgccctgaag gaaattttac atcttgagag acagacggga ggacagctcc tagaagaacc 1980
 taaggctctt cattttaaag gcagcaccca caacctgccc ctgtcaatc acgatatgc 2040
 ccattccctc tggaagagca aattgctggc taaatatcag gaaattccat tttaccatgt 2100
 ttggagtga tctcaaagaa acctgcaact cacttcaact ctggaaagat ttacgctgaa 2160
 25 ccaggtggag ctggtttgca aactctgtgt gcggcaggtg gaaggagaag ggcagatctt 2220
 ccagctcaac tgcacctgtg cagaggaacc tactggcatc gatttgccgc tgctggatcc 2280
 tgcaaacacc atcaccacgg tcacggggcc cagtgtcttc agcatccctc tccctatccg 2340
 gcagaagctc ttagcagcc tggatgcccc ccagacgaga ggccatgact ggaggatgct 2400
 ggcccataag ctgaacctgg acaggtactt gaattacttt gccaccaaat ccagcccaac 2460
 30 tggcgtaatc ctggatcttt gggaagcaca gaacttccca gatggaaaacc tgagcatgct 2520
 ggcagctgtc ttggaagaaa tgggaagaca tgaacgggtg gtgtccttag cagcagaagg 2580
 gcagtattaa ccaccatgct ggaaggggaa at 2612

35 <210> 14
 <211> 636
 <212> PRT
 <213> Homo sapiens

40 <400> 14
 Met Arg Lys Gly Leu Arg Ala Thr Ala Ala Arg Cys Gly Leu Gly Leu
 1 5 10 15
 Gly Tyr Leu Leu Gln Met Leu Val Leu Pro Ala Leu Ala Leu Leu Ser
 20 25 30
 Ala Ser Gly Thr Gly Ser Ala Ala Gln Asp Asp Asp Phe Phe His Glu
 35 40 45
 50 Leu Pro Glu Thr Phe Pro Ser Asp Pro Pro Glu Pro Leu Pro His Phe
 50 55 60
 Leu Ile Glu Pro Glu Glu Ala Tyr Ile Val Lys Asn Lys Pro Val Asn
 65 70 75 80
 55 Leu Tyr Cys Lys Ala Ser Pro Ala Thr Gln Ile Tyr Phe Lys Cys Asn
 85 90 95
 60 Ser Glu Trp Val His Gln Lys Asp His Ile Val Asp Glu Arg Val Asp
 100 105 110
 Glu Thr Ser Gly Leu Ile Val Arg Glu Val Ser Ile Glu Ile Ser Arg
 115 120 125
 65 Gln Gln Val Glu Glu Leu Phe Gly Pro Glu Asp Tyr Trp Cys Gln Cys
 130 135 140
 Val Ala Trp Ser Ser Ala Gly Thr Thr Lys Ser Arg Lys Ala Tyr Val

- 19 -

145 150 155 160
 Arg Ile Ala Tyr Leu Arg Lys Thr Phe Glu Gln Glu Pro Leu Gly Lys
 165 170 175
 5
 Glu Val Ser Leu Glu Gln Glu Val Leu Leu Gln Cys Arg Pro Pro Glu
 180 185 190
 Gly Ile Pro Val Ala Glu Val Glu Trp Leu Lys Asn Glu Asp Ile Ile
 10 195 200 205
 Asp Pro Val Glu Asp Arg Asn Phe Tyr Ile Thr Ile Asp His Asn Leu
 210 215 220
 15 Ile Ile Lys Gln Ala Arg Leu Ser Asp Thr Ala Asn Tyr Thr Cys Val
 225 230 235 240
 Ala Lys Asn Ile Val Ala Lys Arg Lys Ser Thr Thr Ala Thr Val Ile
 245 250 255
 20 Val Tyr Val Asn Gly Gly Trp Ser Thr Trp Thr Glu Trp Ser Val Cys
 260 265 270
 Asn Ser Arg Cys Gly Arg Gly Tyr Gln Lys Arg Thr Arg Thr Cys Thr
 25 275 280 285
 Asn Pro Ala Pro Leu Asn Gly Gly Ala Phe Cys Glu Gly Gln Ser Val
 290 295 300
 30 Gln Lys Ile Ala Cys Thr Thr Leu Cys Pro Val Asp Gly Arg Trp Thr
 305 310 315 320
 Pro Trp Ser Lys Trp Ser Thr Cys Gly Thr Glu Cys Thr His Trp Arg
 325 330 335
 35 Arg Arg Glu Cys Thr Ala Pro Ala Pro Lys Asn Gly Gly Lys Asp Cys
 340 345 350
 Asp Gly Leu Val Leu Gln Ser Lys Asn Cys Thr Asp Gly Leu Cys Met
 40 355 360 365
 Gln Thr Ala Pro Asp Ser Asp Asp Val Ala Leu Tyr Val Gly Ile Val
 370 375 380
 45 Ile Ala Val Ile Val Cys Leu Ala Ile Ser Val Val Val Ala Leu Phe
 385 390 395 400
 Val Tyr Arg Lys Asn His Arg Asp Phe Glu Ser Asp Ile Ile Asp Ser
 405 410 415
 50 Ser Ala Leu Asn Gly Gly Phe Gln Pro Val Asn Ile Lys Ala Ala Arg
 420 425 430
 Gln Asp Leu Leu Ala Val Pro Pro Asp Leu Thr Ser Ala Ala Ala Met
 55 435 440 445
 Tyr Arg Gly Pro Val Tyr Ala Leu His Asp Val Ser Asp Lys Ile Pro
 450 455 460
 60 Met Thr Asn Ser Pro Ile Leu Asp Pro Leu Pro Asn Leu Lys Ile Lys
 465 470 475 480
 Val Tyr Asn Thr Ser Gly Ala Val Thr Pro Gln Asp Asp Leu Ser Glu
 485 490 495
 65 Phe Thr Ser Lys Leu Ser Pro Gln Met Thr Gln Ser Leu Leu Glu Asn
 500 505 510

- 20 -

Glu Ala Leu Ser Leu Lys Asn Gln Ser Leu Ala Arg Gln Thr Asp Pro
 515 520 525
 5 Ser Cys Thr Ala Phe Gly Ser Phe Asn Ser Leu Gly Gly His Leu Ile
 530 535 540
 Val Pro Asn Ser Gly Val Ser Leu Leu Ile Pro Ala Gly Ala Ile Pro
 545 550 555 560
 10 Gln Gly Arg Val Tyr Glu Met Tyr Val Thr Val His Arg Lys Glu Thr
 565 570 575
 Met Arg Pro Pro Met Asp Asp Ser Gln Thr Leu Leu Thr Pro Val Val
 580 585 590
 15 Ser Cys Gly Pro Pro Gly Ala Leu Leu Thr Arg Pro Val Val Leu Thr
 595 600 605
 20 Met His His Cys Ala Asp Pro Asn Thr Glu Asp Trp Lys Ile Leu Leu
 610 615 620
 Leu Ala Gly Val Gln His Pro Ser Leu Leu Ser Gly
 625 630 635
 25
 <210> 15
 <211> 28
 <212> DNA
 <213> Artificial Sequence
 30
 <220>
 <223> Description of Artificial Sequence: Forward Primer
 MF402L11I
 35 <400> 15
 atggagactt cgctcactgg tcctcttt 28
 40
 <210> 16
 <211> 28
 <212> DNA
 <213> Artificial Sequence
 <220>
 45 <223> Description of Artificial Sequence: Reverse Primer
 MF402L11I
 <400> 16
 cctggaggca cataaacct agacatcc 28
 50
 <210> 17
 <211> 21
 <212> DNA
 55 <213> Artificial Sequence
 <220>
 <223> Description of Artificial Sequence: Forward Primer
 MF402L11II
 60
 <400> 17
 acatcctgag gtaggtggtc c 21
 65
 <210> 18
 <211> 21
 <212> DNA
 <213> Artificial Sequence

- 21 -

<220>
<223> Description of Artificial Sequence: Reverse Primer
MF402L11III

5 <400> 18
gggtctcact cattctgttg c 21

10 <210> 19
<211> 21
<212> DNA
<213> Artificial Sequence

15 <220>
<223> Description of Artificial Sequence: Forward Primer
MF402L11IIII

20 <400> 19
ctgtattttg actcaattgt g 21

<210> 20
<211> 21

25 <212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Reverse
Primer MF402L11IIII

30 <400> 20
ccaaaagcct ctccttatta c 21

35 <210> 21
<211> 28
<212> DNA
<213> Artificial Sequence

40 <220>
<223> Description of Artificial Sequence: Forward Primer
MF402L11IV

45 <400> 21
gaatacagac acacagagac agcacctt 28

<210> 22
<211> 28
<212> DNA
<213> Artificial Sequence

50 <220>
<223> Description of Artificial Sequence: Reverse Primer
MF402L11IV

<400> 22
gttccttgct ccagtgaag acataatc 28

60 <210> 23
<211> 28
<212> DNA

65 <213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Forward Primer

- 22 -

MF402L11V

5 <400> 23
ctaacaccag ttcagtggaa cccattgc 28

10 <210> 24
<211> 28
<212> DNA
<213> Artificial Sequence

15 <220>
<223> Description of Artificial Sequence: Reverse Primer
MF402L11V

20 <400> 24
atttgatgtt catacagggt tgggggac 28

25 <210> 25
<211> 28
<212> DNA
<213> Artificial Sequence

30 <220>
<223> Description of Artificial Sequence: Forward Primer
MF402L11VI

35 <400> 25
ctccacttgg actctgatga caccactg 28

40 <210> 26
<211> 27
<212> DNA
<213> Artificial Sequence

45 <220>
<223> Description of Artificial Sequence: Reverse Primer
MF402L11VI

50 <400> 26
catagttcac tgcagcctca acctcct 27

55 <210> 27
<211> 28
<212> DNA
<213> Artificial Sequence

60 <220>
<223> Description of Artificial Sequence: Forward Primer
MF90A19I

65 <400> 27
taattatgat tagacagacc gtcctcca 28

70 <210> 28
<211> 28
<212> DNA
<213> Artificial Sequence

75 <220>
<223> Description of Artificial Sequence: Reverse Primer
MF90A19I

80 <400> 28

20 <210> 30
 <211> 28
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Description of Artificial Sequence: Reverse Primer
 MF90A19II
 25 <400> 30
 gatggtgtgg aggaagaggt tcaaaggt 28
 30 <210> 31
 <211> 28
 <212> DNA
 <213> Artificial Sequence
 35 <220>
 <223> Description of Artificial Sequence: Forward Primer
 MF90A19III
 40 <400> 31
 aactgagaac tggaaggtag ctttggga 28
 45 <210> 32
 <211> 28
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Description of Artificial Sequence: Reverse Primer
 MF90A19III
 50 <400> 32
 acaagagcag tcagaactat ttggccct 28
 55 <210> 33
 <211> 28
 <212> DNA
 <213> Artificial Sequence
 60 <220>
 <223> Description of Artificial Sequence: Forward Primer
 MF90A19IV
 65 <400> 33
 atttgctggc tccgacagta agtattcc 28

- 24 -

<210> 34
<211> 28
<212> DNA
<213> Artificial Sequence
5
<220>
<223> Description of Artificial Sequence: Reverse Primer
MF90A19IV
10 <400> 34
atgcagtgat gcttaacacc ctgtctct 28
15 <210> 35
<211> 28
<212> DNA
<213> Artificial Sequence
<220>
20 <223> Description of Artificial Sequence: Forward Primer
MF90A19V
<400> 35
cctgctctgt ctgccttgga aaagaaat 28
25
<210> 36
<211> 28
<212> DNA
30 <213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Reverse Primer
MF90A19V
35
<400> 36
tcttgGCCa gacggtataa gagtccat 28
40
<210> 37
<211> 28
<212> DNA
<213> Artificial Sequence
45 <220>
<223> Description of Artificial Sequence: Forward Primer
MF90A19VI
<400> 37
50 agtgcaccag taagcatgcc ttccaata 28
55
<210> 38
<211> 28
<212> DNA
<213> Artificial Sequence
<220>
60 <223> Description of Artificial Sequence: Reverse Primer
MF90A19VI
<400> 38
aagctggcaa ggagattctc aaaagtgg 28
65
<210> 39
<211> 28
<212> DNA

- 25 -

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Forward Primer
MF90A19VII

<400> 39

accacatctg ctgaagttga ctggacct

28

10

<210> 40

<211> 28

<212> DNA

<213> Artificial Sequence

15

<220>

<223> Description of Artificial Sequence: Reverse Primer
MF90A19VII

20

<400> 40

tctgggcaac atagtgaac cctgtctc

28

25

<210> 41

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

30 <223> Description of Artificial Sequence: Forward Primer
MF90A19VIII

<400> 41

cttgaacctg ggagacggag gttgcagt

28

35

<210> 42

<211> 28

<212> DNA

40 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Reverse Primer
MF90A19VIII

45

<400> 42

aagcatatct ctgaccctct ggccattc

28

50

<210> 43

<211> 28

<212> DNA

<213> Artificial Sequence

55

<220>

<223> Description of Artificial Sequence: Forward Primer
MF90A19IX

<400> 43

60 gaagctgcaa tcccatcctt gtctcttt

28

<210> 44

<211> 28

65 <212> DNA

<213> Artificial Sequence

<220>

- 26 -

<223> Description of Artificial Sequence: Reverse Primer
MF90A19IX

<400> 44
5 tgggtggtgca cacctgtaat accagcta 28

<210> 45
<211> 20
10 <212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Forward Primer
15 C4P3F

<400> 45
aggatgtcca tggtttttggga 20

20 <210> 46
<211> 20
<212> DNA
<213> Artificial Sequence

25 <220>
<223> Description of Artificial Sequence: Reverse Primer
C4P3R

30 <400> 46
ctgtttccag cttccagagg 20

<210> 47
35 <211> 20
<212> DNA
<213> Artificial Sequence

<220>
40 <223> Description of Artificial Sequence: Forward Primer
C4P4F

<400> 47
45 catggtccca aagctctgtt 20

<210> 48
<211> 20
<212> DNA
50 <213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Reverse Primer
C4P4R

55 <400> 48
tgacggaatt tcagtgcaac 20

60 <210> 49
<211> 20
<212> DNA
<213> Artificial Sequence

65 <220>
<223> Description of Artificial Sequence: Forward Primer
C4P7F

- 27 -

<400> 49
tctgatacag gggatttggc 20

5 <210> 50
<211> 20
<212> DNA
<213> Artificial Sequence

10 <220>
<223> Description of Artificial Sequence: Reverse Primer
C4P7R

15 <400> 50
ttcaggtggg tgatgtgtgt 20

20 <210> 51
<211> 20
<212> DNA
<213> Artificial Sequence

25 <220>
<223> Description of Artificial Sequence: Forward Primer
C4P8F

30 <400> 51
aggaaagagg ggtgctgtct 20

35 <210> 52
<211> 20
<212> DNA
<213> Artificial Sequence

40 <220>
<223> Description of Artificial Sequence: Reverse Primer
C4P8R

45 <400> 52
atagcaggag agtgagggca 20

50 <210> 53
<211> 20
<212> DNA
<213> Artificial Sequence

55 <220>
<223> Description of Artificial Sequence: Forward Primer
C4P9F

60 <400> 53
ggccacttt ctggttcata 20

65 <210> 54
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Reverse Primer
C4P9R

<400> 54
tcagccagac cagtgtgaaa 20

- 28 -

<210> 55
<211> 20
<212> DNA
5 <213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Forward Primer
C15P1F
10
<400> 55
ggctctgtgcc ctctgtcaag 20

15 <210> 56
<211> 20
<212> DNA
<213> Artificial Sequence

20 <220>
<223> Description of Artificial Sequence: Reverse Primer
C15P1R

<400> 56
25 aggccttccca cacagatgag 20

<210> 57
<211> 20
30 <212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Forward Primer
C15P2F
35
<400> 57
cactcttggc acccttaacc 20

40
<210> 58
<211> 20
<212> DNA
<213> Artificial Sequence

45 <220>
<223> Description of Artificial Sequence: Reverse Primer
C15P2R

50 <400> 58
tttggtctgt cctccccaac 20

55 <210> 59
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
60 <223> Description of Artificial Sequence: Forward Primer
C15P5F

<400> 59
65 tttcatcaaa gcacccaca 20

<210> 60
<211> 20

- 29 -

<212> DNA
<213> Artificial Sequence

5 <220>
<223> Description of Artificial Sequence: Reverse Primer
C15P5R

10 <400> 60
tggttaattcg tgctgcagag 20

<210> 61
<211> 20
<212> DNA
15 <213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Forward Primer
C15P6F

20 <400> 61
cccaaacatt ccaacttcgt 20

25 <210> 62
<211> 20
<212> DNA
<213> Artificial Sequence

30 <220>
<223> Description of Artificial Sequence: Reverse Primer
C15P6R

35 <400> 62
tagctcaagc ctgaacccat 20

<210> 63
<211> 26
40 <212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Forward Primer
BPC4F

45 <400> 63
accctaattgc ccatttttgg ttgagg 26

50 <210> 64
<211> 28
<212> DNA
<213> Artificial Sequence

55 <220>
<223> Description of Artificial Sequence: Reverse Primer
BPC15R

60 <400> 64
aaataaccagt cctttccgaa gcctctcc 28

65 <210> 65
<211> 28
<212> DNA
<213> Artificial Sequence

- 30 -

<220>
<223> Description of Artificial Sequence: Forward Primer
BPC15F

5 <400> 65
cgacagaacc atccaaacca agacactt 28

10 <210> 66
<211> 28
<212> DNA
<213> Artificial Sequence

15 <220>
<223> Description of Artificial Sequence: Reverse Primer
BPC4R

20 <400> 66
gtgtttattg tgtcacgtgc agttgctg 28

<210> 67
<211> 27
<212> DNA

25 <213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Forward Primer
RORa1 5' UTR

30 <400> 67
ctgaaaacag aagatagagg gagtctc 27

35 <210> 68
<211> 20
<212> DNA
<213> Artificial Sequence

40 <220>
<223> Description of Artificial Sequence: Reverse Primer
UNC5c 3'UTR

45 <400> 68
atttcccctt ccagcatggt 20

<210> 69
<211> 27

50 <212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Forward Primer
AP1

55 <400> 69
ccatcctaatac gactcact atagggc 27

60 <210> 70
<211> 22
<212> DNA
<213> Artificial Sequence

65 <220>
<223> Description of Artificial Sequence: Reverse Primer

- 31 -

<400> 70
ctccacagat cttgcatgga at 22

5 <210> 71
<211> 23
<212> DNA
<213> Artificial Sequence

10 <220>
<223> Description of Artificial Sequence: Forward Primer
AP2

15 <400> 71
actcactata gggctcgagc ggc 23

20 <210> 72
<211> 23
<212> DNA
<213> Artificial Sequence

25 <220>
<223> Description of Artificial Sequence: Forward Primer

<400> 72
tctgtgtttc ataccagga ggt 23

30 <210> 73
<211> 23
<212> DNA
<213> Artificial Sequence

35 <220>
<223> Description of Artificial Sequence: Reverse Primer

40 <400> 73
agaacacaca cccatagact gca 23

45 <210> 74
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Forward Primer

50 <400> 74
cctgtaccgc atttggcag 19

55 <210> 75
<211> 20
<212> DNA
<213> Artificial Sequence

60 <220>
<223> Description of Artificial Sequence: Reverse Primer

<400> 75
tgcaaaccag ctccactgtg 20

65